Summer 1998

# Analyzing the Effects of Display Characteristics and Cognitive Variables on Performance Using Keystroke and Eye Movement Data

Orhan E. Beckman
*Old Dominion University*

Recommended Citation
Beckman, Orhan E.. "Analyzing the Effects of Display Characteristics and Cognitive Variables on Performance Using Keystroke and Eye Movement Data" (1998). Doctor of Philosophy (PhD), dissertation, Psychology, Old Dominion University, DOI: 10.25777/0gy4-d116
https://digitalcommons.odu.edu/psychology_etds/144

# ANALYZING THE EFFECTS OF DISPLAY CHARACTERISTICS AND COGNITIVE VARIABLES ON PERFORMANCE USING KEYSTROKE AND EYE MOVEMENT DATA

by

Orhan E. Beckman
B.A. June 1990, Earlham College
M.S. May 1994, Old Dominion University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

INDUSTRIAL/ORGANIZATIONAL PSYCHOLOGY

OLD DOMINION UNIVERSITY
August 1998

Approved by:

Glynn D. Coates (Director)

James R. Comstock, Jr. (Member)

Terry L. Dickinson (Member)

Randall L. Harris, Sr. (Member)

# ABSTRACT

## ANALYZING THE EFFECTS OF DISPLAY CHARACTERISTICS AND COGNITIVE VARIABLES ON PERFORMANCE USING KEYSTROKE AND EYE MOVEMENT DATA

Orhan E. Beckman
Old Dominion University, 1998
Director: Dr. Glynn D. Coates

Information about how operators use their eyes while interacting with visual displays is often an overlooked aspect of human-computer interaction. Such information is fundamental to assessing the quality of software interfaces and understanding the cognitive processes that underlie operator behavior. Other research evaluating information displays evolved from using reaction time and subjective data as dependent variables to using oculometric measures. In the current research conventional performance measures are coupled with oculometric measures to evaluate the influence display characteristics and cognitive variables have on performance.

Twelve subjects used a software program to complete a series of specified tasks. Subjects were asked to search for 36 items from the database in a serial manner. Both keystroke and oculometric data were recorded while the subjects used the software database. Four dependent variables were derived from this data: task time, error rate, dwell time and dwell frequency. The four independent variables were information density, display layout, task complexity, and experience.

Out of the four independent variables used in the current research, task complexity, a cognitive variable, clearly had the largest effect on both the time-based measures of performance and the oculometric measures of performance. Task complexity yielded a main effect in the task time data, the error rate data, the dwell time data and the dwell frequency data. Increases in task complexity yielded increases in task time, error rate, dwell time and dwell frequency. The results also showed that local information density had an effect on task time but only when overall density of the software interface was higher. While it was found that information density had a consistent effect on the frequency of dwells these results support other research that shows information density has a limited effect on performance. The display layout variable also had a limited influence on both performance and oculometric measures.

# ACKNOWLEDGEMENTS

I would like to thank Dr. Glynn Coates whose support and guidance over the years has allowed me to excel in my graduate studies at Old Dominion University and in the application of my skills working at Hewlett-Packard. The impact Dr. Coates has had on the quality of this work and the quality of my professional development cannot be overstated. I would also like to thank Dr. Ray Comstock, Jr., Dr. Terry Dickinson and Dr. Randall Harris, Sr. for their support and guidance.

In addition, I would also like to thank Tom Splaver of the VPR Systems Lab at Hewlett-Packard who provided the support and encouragement for me to finish my Doctor of Philosophy degree. I am indebted to Hewlett-Packard for the generous funding I was granted through the Resident Fellowship Program which made it possible for me to finish my studies at Old Dominion University.

Finally, I owe my deepest appreciation to my wife Marcia who unconditionally supported and tolerated me in my slow but steady pursuit of my dream to finish.

# TABLE OF CONTENTS

Page

# LIST OF FIGURES

FIGURE Page

# INTRODUCTION

## Overview

Information about how operators use their eyes while interacting with visual displays is often an overlooked aspect of human-computer interaction. Such information is fundamental to assessing the quality of software interfaces and understanding the cognitive processes that underlie operator behavior. Research evaluating information displays evolved from using reaction time (Graham, 1956) and subjective data as dependent variables to using oculometric measures (Harris & Christhilf, 1980). Conventional behavioral indices of software interface quality consist of time-based measures, error rates, and subjective measures. Software interface research and development can benefit from an evolution in dependent measures.

The visual display terminal of a computer workstation is the main source of feedback to its operator. Spatially oriented software interfaces have replaced symbolic displays and are now standard in the computer industry. In this new graphic user interface paradigm, tasks can be performed directly on spatial arrays rather than negotiating abstract symbols (Shneiderman, 1987). This spatial metaphor makes visual search an integral component of human-computer interaction. The study of eye scan patterns can increase our understanding of the information flow characteristics between the computer and its operator when other dependent measures yield little or no data (Graf & Krueger, 1989; Moray & Rotenberg, 1989) or when conventional measures of performance lack adequate resolution (Dumas & Redish, 1994).

In the current research, four independent variables are manipulated and their

---

This dissertation adheres to the format of the Publications Manual of the American Psychological Association (1996).

effects are assessed using both conventional measures of performance and measures of ocular behavior. Two independent variables, information density and interface layout, are related to display characteristics. The two other variables, task complexity and experience, are cognitive in nature. The conventional measures of performance include time and errors. The eye movement measures are derived from oculometric data. Four analyses were performed. Two analyses use conventional measures of performance. The second two analyses utilize eye movement measures. Results of the analyses will be compared. The incremental validity of eye movement data, or the amount of information eye movement data yield beyond that which is provided by the conventional measures, will be assessed.

In short, this research should help to answer the following questions. One, what affect do these display characteristics and cognitive variables have on performance? Two, how can eye movement data be used to understand differences in performance due to manipulation of the independent variables? Finally, what unique and useful information do eye movement data offer beyond that which is available through conventional measures of performance? Next, a brief review of the research that has been conducted on scan patterns and information displays is provided.

Fitts, Jones, Milton and Cole (1950) conducted the first definitive research examining the natural scanning patterns of pilots. They monitored the eye movements of Air Force pilots as they performed flight maneuvers in order to facilitate more efficient pilot training and aircraft instrument panel layout. Link analysis was used as a method for analyzing eye movement data (Sanders & McCormick, 1993). Sequential link values of eye movements between the instruments were derived from the data. Other parameters obtained from the eye movement data included average fixation length, fixation rate, and the percentage of time spent viewing each instrument. The authors concluded that frequency of eye fixations, or what is referred to in the current research as dwells, is a reflection of the importance of the object being fixated. The length of the fixation was used to assess the difficulty of the interpretation. The spatial arrangement of the instruments on the display influenced the pattern of eye movements. The patterns, derived from the link analyses, were considered to reflect the goodness of spatial arrangement of the displays. The pilot's task and experience level were also found to contribute systematic variance to the eye movement patterns.

Fitts et al. (1950) described concisely the usefulness of the eye movement measure when he wrote, "If we know where a pilot is looking we do not necessarily know what he is thinking, but we know something of what he is thinking about" (p. 24). The results of this research were used to design the basic "T" arrangement of instruments in a cockpit that is still a standard today. This research provided a benchmark for later oculometer studies (Donk, 1994; Senders, 1966). Since the work of Fitts et al. (1950), eye movements have been used in research examining the scanning patterns of pilots (Christhilf, 1980; Harris &; Spady, 1978; Jones, 1985), radiologists (Gale & Worthington, 1984; Kundel, Nodine, Toto, 1984), television viewers (Flagg, 1978) and industrial inspectors (Drury, 1975).

Eye movements are a product of both environmental and internal, or cognitive, factors (Harris & Spady; 1985; Wickens 1992). The relative amount of influence these variables have on ocular behavior is debated in the scientific literature (Tullis, 1983; Wickens, 1987). The four independent variables used in this research can be classified into a display-oriented group (information density and display arrangement) and a cognitive-oriented group (experience and task complexity). Research relating to these variables is reviewed below.

## Information Density

Information density, a display characteristic, is cited in the literature as a factor affecting scan patterns and visual search time. Tannas (1985) considers information density one of the most important characteristics of any visual display. Holahan, Culler and Wilcox (1978) demonstrated a positive relationship between the level of visual distraction in a display space and reaction time. Their research showed that the ability to locate and respond to a stop sign in a cluttered display was directly inhibited by the proximity of other irrelevant signs in the field of view. Landis, Slivka & Jones (1967) proposed that the general function relating quality of performance and display density has an inverted-U shape. At low levels of density, raising the density enhances performance while at high levels it inhibits performance. This implies there may be an optimal level.

Tullis (1983) identified four information density characteristics of alphanumeric displays, overall density, local density, grouping and layout complexity, and found that these characteristics correlate with search time and eye movement parameters. Kollers, Duchinsky & Ferguson (1981) compared single spaced with double spaced displays of text on a cathode-ray tube. Single spacing required more eye fixations per line, resulted in fewer words read per fixations and required longer total reading time. Research suggests that the lower search times associated with icons versus words may be partially a function of differences in information density (Lansdale, Jones & Jones, 1989). Text involves a larger number of lines close together that are more difficult to resolve in peripheral vision.

Scott (1992) used spatial frequency grids to examine what influence cycle frequency, high/low contrast, and high/low similarities of non-targets have on search time. While all variables affected search time, spatial frequency of cycles had a pronounced effect on search time, thirteen times greater than that of contrast. Scott also recorded eye movements during the task and found that if the target was detectable in peripheral vision, fewer fixations were produced before the target was located and the search was less systematic as indicated by a transition matrix. The results of these studies suggest that information density influences both performance times and visual information acquisition patterns.

Other researchers argue that information density has little effect on performance time. High-density environments retard performance a little but also require less visual scanning, with more information captured per fixation. Lower display density results in greater scanning distances but less performance attenuating clutter. Thus the two factors, visual scanning and visual clutter, essentially trade off with one another as target dispersion changes. Wickens & Andre (1990) found the most critical variable in predicting performance is the degree of separation of relevant from irrelevant items and not the density of relevant items themselves. Although information density guidelines exist, no one has manipulated these characteristics over a wide enough range to validate either camp's assertions. The researchers all appear to agree that while information density may or may not affect time-based measures of performance, it does influence ocular behavior.

## Spatial Layout

The layout of instruments in physical space was shown by Fitts et al. (1950) to influence eye movements. The goodness of different instrument configurations was assessed using the link values in a transition matrix. Both Senders (1983), who proposes a normative model of visual sampling and Van Delft (1987), who advocates sequential sampling heuristics as determinants of scan patterns, do not predict any dependency of sampling on instrument arrangement. Donk (1994) in a test of Senders' (1983) normative model of visual sampling behavior reported spatial arrangement as one of two major sources of variance in visual sampling behavior.

In accordance with the normative model, Donk (1994) found sampling behavior was determined in part by the information generation rates of the four instruments that constituted the display in his study. Scan behavior was also strongly affected by the spatial arrangements of the instruments, with horizontal transitions occurring more often and diagonal transitions less often than would be predicted by the normative model. Display configuration has since been cited in other research as a variable influencing scan patterns (Kolers et al. 1981) and response time (Treisman, 1982; Tullis, 1983).

Others contend that cognitive factors play a greater role in determining scan patterns (Levy-Schoen, 1981; Wickens, 1992). These researchers argue that location driven search tendencies are not strong and scan strategies are dominated primarily by cognitive factors. More research is needed to understand the influence display characteristics, such as information density and spatial arrangement, have on performance during the use of visual interfaces.

## Information

Regions that yield high amounts of information disproportionately attract eye fixations (Mackworth & Morandi, 1967; Yarbus, 1967). According to Senders (1983), information theory (Shannon, 1948) dictates that the sampling frequency of an instrument yielding status, such as a cockpit display, is related to its bandwidth. In Senders' (1983) normative model, which uses concepts derived from information theory, monitoring performance is described as a direct function of the information generation rates of the stimuli. In this model, the frequency of sampling is a linear function of the instrument's rate of information generation. This implies that the frequency of eye fixations on a display will increase as the amount of information the instrument produces increases.

Dwell time, what Fitts described as "average fixation length", has been used as a measure of importance and information content (Fitts et. al, 1950). Harris and Christhilf (1980) found that visual dwell times were short (< (0.5) second) when pilots monitored an instrument to see if a needle was at its expected level. When the display's information content was higher, reflecting a change in an underlying state of the system, the authors found that fixations were considerably longer (>= 1.0 second). Wickens (1992) suggests that dwell length and the amount of information extracted are correlated but not perfectly.

Low familiarity, low frequency, and out of context information translate directly to higher information content. Fixation dwells are also related to the difficulty of information extraction. Displays that are less legible or contain higher amounts of information will result in longer fixations. Information transmission can be thought of as a relation between the subject and scene rather than simply a property of the visual stimuli itself. The operator's eyes are attracted to areas that have the highest probability of reducing the error signal inputs he/she receives.

## Cognitive Processes

Eye movement data provide a rich source of information in support of cognitive processes, as evidenced by two decades of research demonstrating a relationship between cognitive processes and eye movement (Just & Carpenter, 1976). Cognition, including attention, expectations and strategies, is considered an important factor in determining scan paths and fixations (Boff, Kaufman & Thomas, 1987).

Using fixed ambiguous visual stimuli, Stark & Ellis (1981) demonstrated how changes in cognitive state can influence scan patterns. Stark & Ellis (1981) measured eye movements before and after identification of an object in an ambiguous stimulus. By holding all variables constant other than cognition, this research demonstrates the influence cognition can have on eye movements. Zero, first and second-order Markov matrices were used to quantify the scan patterns. Such matrices were used to identify sequential strings of successive fixation points from which the underlying cognitive processes that control the eye movements can be better understood. Markov models are a particularly interesting analysis technique because subjects are usually not aware of the patterns of their eye movements and yet these movements reflect the dynamics of the system in which they interact. Other quantitative analysis methods that can be used to assess underlying cognitive structures are described in the literature (Harris, Glover & Latimer, 1988; Scinto & Barnette 1986; Spady, 1986).

Wickens (1987) defines a mental model as “a set of expectancies about how frequently and when events will occur on each channel and about the correlation between events on pairs of channels” (p. 527). Since a cognitive model is a mental construct, it is not directly observable or measurable. Inferring a mental model from indirect methods, such as reaction time or error data, is difficult. Eye movements can, at the very least, be considered tags or experimentally accessible quantities that researchers can observe to understand underlying processes of cognition (Stark & Ellis, 1981). Since scan patterns reflect changes in cognitive state, a stronger hypothesis is that mental models direct scan path movements. The ability to record eye movement provides a structured way to understand externalized aspects of information processing arising from mental models. Object hypothesis advocates argue that lower order aspects of physical stimuli generally determine eye movements (Didday & Arbib, 1972; Michels and Zusne, 1965). The influence high-order cognitive factors, such as an understanding of the system under focus or the information being sought, have on scan patterns is documented in the scientific literature (Donk, 1994; Senders, Elkind, Grignetti & Smallwood, 1964; Stark & Ellis, 1981; Yarbus, 1967).

If scanning behavior reflects the operator's mental model of the environment, it can be used to indicate his or her information needs. This has important implications for increasing our ability to measure operator strategies when interacting with information sources. Krappman (1995) studied the eye movements of subjects directing a computer-simulated factory. The subjects had no experience with the simulation. The strategies employed by the subjects could be inferred from the scan patterns they exhibited and Krappman could, post hoc, differentiate successful subjects from unsuccessful subjects based on this criterion. Others found that in the first trial of a complex problem-solving situation, fixation frequency data could be used to infer information gathering strategies and discriminate successful subjects from unsuccessful ones (Luer, Hubner & Lass, 1985; Luer, Lass, Ulrich & Schroiff, 1986; Russo and Rosen, 1975).

Process control researchers use eye movements as means to trace the progress of information processing during periods of 'inactivity' or 'cognitive lockup' (Moray &

Rotenberg, 1989). Eye movement is an appropriate dependent measure for process control research where skilled operators spend long periods just observing the system with only occasional interventions. Cognitive lockup, in the context of fault management, is a tendency for the subject to ignore parts of the system because of hypotheses generated about the state of the system. Moray & Rotenberg (1989) found eye movement data could reveal detailed information about the information processing patterns of operators during periods of inactivity. Cognitive lockup was found to be a result of the serial treatment of faults. The treatment of problems followed the order of occurrence, independent of the problem's severity. This research highlights the unique information and insights eye movement data can offer the researcher when conventional dependent measures yield little data or, in the case of think aloud protocols, influence the construct being measured (Harris & Spady, 1985; Lass, Klettke, Luer & Ruhlender, 1991; Wright & Converse, 1992).

Keystroke data has been heralded by many in the usability engineering field as a valuable performance measure in software based tasks. Critics argue that keystroke data alone provides a detailed but limited record of interaction with software (Dumas & Redish, 1994). The current author argues that keystroke data supplemented with oculometric data can provide a significantly richer account of behavioral interaction with a software interface than keystroke data can alone. This provides a base from which cognitive structures underlying performance can be more easily measured and understood. Others in the field agree (Scott, 1991; Scott-Findlay, 1989; Faraday & Sutcliffe, 1997).

## Experience

Research has repeatedly demonstrated that experience influences the patterns of eye movements. Fitts et al. (1950) reported that more experienced pilots exhibited a tendency to make shorter fixations on instruments than less experienced pilots. Demaio, Parkinson, Leshowitz, Crosby and Thorpe (1976) found less-experienced pilots exhibited considerably more statistical dependency in their scan patterns, than seasoned pilots. This was interpreted to be a reflection of more conscious shifts in attention by the less-experienced pilots. Harris, Tole, Stephens and Eprath (1982) found an operator's skill level in a man-machine control task affects both temporal and sequential aspects of scan patterns. Others share similar conclusions regarding the effects of interface familiarity on scanning behavior (Graf & Krueger, 1989; Stark & Ellis, 1981).

Target search, an inherent behavioral component of spatially oriented interfaces, is considered by most in the field to be driven in part by cognitive factors. These factors relate to the expectancy of where in the display a target containing the most useful information is likely to be found. These areas tend to be fixated first and most frequently. Such patterns of information-seeking and scanning behavior have been used to account for differences between novices and experts (Abernathy, 1988). Areas of high information in the visual field attract fixations. Scan paths over same visual stimuli will vary according to changes in experience, goals and expectations. Information transmission is therefore not a static property but varies in accordance with situational characteristics.

Training of efficient oculomotor strategies has been shown to improve performance (Jones, 1985; Spady, Jones, Coates & Kirby, 1982). Although simply viewing efficient scan patterns has proven an effective means for improving behavior, showing students their own scan patterns and having them actively participate when viewing optimal scan path behaviors can more rapidly shape performance (Shapiro &

Raymond, 1989).

## Modern Visual Displays

Information displays both inside and outside the cockpit have evolved since the 1950's (Williams & Harris, 1985). Tullis (1983) wrote, the "number of displays in use and number of people working on them is overwhelming" (p. 658). His statement is as valid today as it was fifteen years ago. While the number of information displays in our technological society and the quantity of information available through information displays has increased, behaviorally based measures of the quality of visual interface designs have not been similarly expanded (Scott, 1991).

The visual interface of the modern personal computer is spatially oriented as opposed to symbolic. Sutherland (1963), in his Ph.D. dissertation, first demonstrated this direct manipulation-style of human-computer interaction in the Sketchpad system. The philosophy behind the Sketchpad system is that the computer should be manipulated in much the same way objects in the real world are manipulated. Today direct manipulation interfaces are filled with familiar objects such as windows, folders and buttons.

Hypertext, a term coined by Nelson (1965), is now a widely used interface convention due, in large part, to the popularization of the graphical portion of the Internet called the World Wide Web. The idea for hypertext is usually credited to Bush (1945) for his MEMEX idea. Manipulation of objects, activation of hypertext functions and navigation in display space is achieved primarily through the use of a spatial input device such as a mouse. The symbolic software interfaces of yesterday emphasized syntactic structure. Spatially oriented software interfaces emphasize appearance and location. The capitalization on the spatial metaphor has changed the manner in which people interact with computers. Visual search is now an integral component of human-computer interaction. The operator is more likely to search for information on the screen than search his long term memory for syntax to enter at a command line (Schniederman, 1983). Another implication of this change is a reduced reliance on the keyboard for input and navigation, and an increased reliance on spatial input devices.

The direct manipulation interface is now standard in the personal computer industry. The effects spatial interface conventions have on visual behavior and search strategies are debated. Some research has focussed on the relationship of eye movements and pull-down menu use (Giroux & Belleau, 1986; Lee & MacGregor, 1985). Hendrickson (1988) found that visual performance varied as a function of window size, the number of active windows and query length, a cognitive variable. Hendrickson demonstrated the influence both display and cognitive characteristics can have on visual performance in human-computer interaction. Displaying status information at the mouse cursor, or point of regard, has been shown to increase overall performance times (Scott & Findlay, 1991). These results coincide with Russo's (1978) contention that eye movements involved in search are likely to exact a cognitive cost and thereby increase response time. Research focussing on the visual characteristics of icons that facilitate the visual search task has yielded different conclusions regarding serial and parallel search patterns (Lansdale, Jones & Jones, 1989; Scott & Findlay 1991; Treisman & Souther, 1985). The need to increase our understanding about how computer users visually scan for information, what strategies users employ, and what effects software interface conventions have on visual performance has been highlighted by the researchers in the eye movement field (Graf & Krueger 1989; Scott & Findlay, 1991).

Despite exponential increases in the number of displays in our society and their widespread use, computer users often experience frustration in accessing and interpreting the information from visual display terminals. As the amount of digital information available through the average display terminal increases, aspects of visual displays such as the ability to navigate through the interface and the ease with which information can be extracted from the display, become increasingly important (Tullis, 1983).

The current research is designed to further our understanding of the effect information density and spatial layout, two display characteristics, and experience and task complexity, two cognitive variables, have on human performance. This research is also designed to assess the utility of eye movement data in the analysis and understanding of human-computer interaction. As evidenced in the literature review above, information density and interface layout have been shown to influence both performance and scan patterns, although the effects are not well understood. The effects the two display characteristics and the two cognitive attributes have on behavior are analyzed using both conventional and oculometric measures. The effects of information density and display layout are analyzed first using conventional performance measures. These performance measures include time on task and errors. Second, the influences these variables have on eye movements are analyzed using dependent measures derived from oculometric data. Through this process, the effect information density, spatial layout, experience, task complexity have on performance will be clarified and the utility of using eye movement data in the analysis of performance will be assessed.

## Design Overview

Four independent variables, information density, interface layout, experience with a software graphical user interface and task complexity, are used in the current research. Two types of observer responses, eye movement and mouse input, were recorded. Keystroke data via the mouse are used to identify operator inputs, derive time-based performance measures, and allow the eye movement data to be synchronized with changes in the software interface.

Scanning behavior is a very complex phenomenon (Harris, Glover & Spady, 1986). Because eye movement data may be analyzed in many different ways, it is important that all constructs have operational definitions (Comstock, 1983; Harris et. al, 1986). A lookpoint is current X (horizontal) and Y (vertical) coordinates in the visual field indicating where on the specified plane the subject is looking at any one thirtieth of a second. Dwell time, or a dwell, is the time spent looking at an instrument or display object.

Related to dwell time, fixation time is defined as a series of lookpoints that do not exceed a selected boundary limit. Because of measurement errors, a radius, not a specified point, is used to define fixations. The radius used to define the area around a lookpoint within which the next lookpoint must fall to be considered part of or contributing to a fixation is 35 units out of 511 vertical units and 511 horizontal units. The target plane is defined using the arbitrary index of 511 units horizontally and 511 units vertically. Multiple fixations are possible within a single dwell and movement from fixation point to fixation point can be within or between display areas or objects. From these base definitions other dependent measures such as average dwell time, dwell percentage, fixation frequency, transition matrices, and transition rate can be derived. These other measures are defined below as they are used.

Data being collected by the oculometer reflect one of three basic states. The first is an out-of-track condition. In this condition the oculometer cannot determine where the subject is looking, such as during a blink, rapid head movement or when the subject's lookpoint is outside the bounds of the specified plane. The byte of data indicating the plane under focus equals zero when the subject's lookpoint is out of track. When it is in track this byte equals one. The second condition is transition. In a transition a lookpoint is not part of or is not forming a new fixation. In the third and final possible state, fixation, a lookpoint, or a series of lookpoints, is starting a new fixation or contributing to an existing fixation being within the specified area around the previous lookpoint.

## Hypotheses

The following hypotheses are tested in the present research.

Hypothesis A: Displays with high or low information density are more difficult to interpret and result in lower rates of information transfer than display areas with moderate information density.

Hypothesis A is tested using task time and dwell time as dependent measures. Areas with high or low information density contribute to longer task completion times than display areas with moderate information density. Dwell Time: Areas with high or low information density contribute to longer dwell times than display areas with moderate information density.

Hypothesis B: The arrangement of display elements in a visual field influences performance.

Hypothesis B: is tested using task time and dwell time as dependent measures. The arrangement of display elements in the software hierarchy will influence the time required to complete the task. The arrangement of display elements in the software hierarchy will influence the dwell time required to complete the tasks.

Hypothesis C: The location of display elements will influence strategies employed by the subjects.

Hypothesis C will be tested using the dwell frequency as a dependent measure. The location of display elements will influence strategies as reflected in the frequency of dwells on different display elements.

Hypothesis D: Experience influences strategy.

Hypothesis D will be tested using task time and dwell frequency as dependent measures. The strategies adopted by the subjects will be reflected in the task time data. The strategies adopted by the subjects will be reflected in the dwell frequency data.

# METHOD

In this research, twelve subjects used a software program to complete a series of specified tasks. The software program was a custom database of computer peripherals. The database program presented a graphical user interface that consisted of buttons and hypertext. Subjects were asked to search for 36 items from the database in a serial manner. Both keystroke and oculometric data were recorded while the subjects interacted with the software database. Four dependent variables were derived from this data: task time, error rate, dwell time and dwell frequency. Four independent variables--information density, display layout, task complexity, and experience--were used. The test period took between six to ten minutes for each subject.

This study uses a 3 x 2 x 3 x 2 mixed-model design. The design combines the information density condition, display layout condition, task complexity condition and the experience condition factorially. The information density (high, medium, low), display layout (present or absent navigation aid) and task complexity condition (high, medium, low) are used as within-subject variables and the experience condition (high or low) was the sole between-subject variable.

There are four main conditions. In the first condition, display layout, subjects used the database program with or without a navigation aid. In the navigation aid present condition, information about the contents of the software database is always displayed in the left portion of the visual display (see Figure 1).

The navigation aid provided shortcuts to other areas of the software database program. Once a subject entered a particular area of the database by clicking on the corresponding button the difference between the navigation aid present and navigation aid

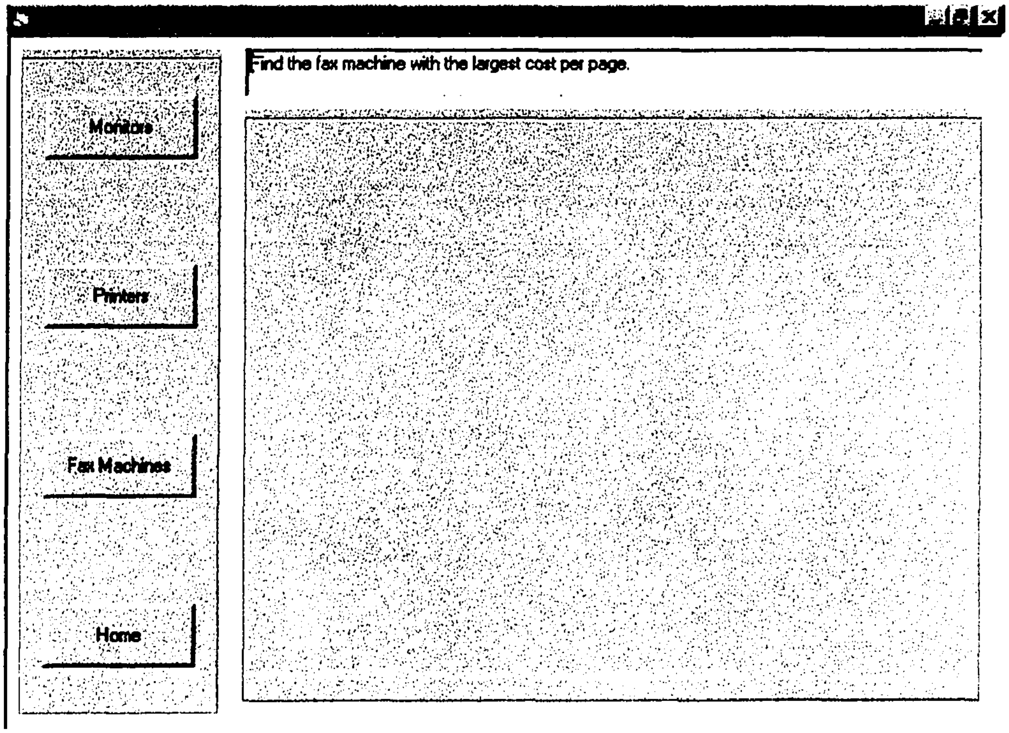
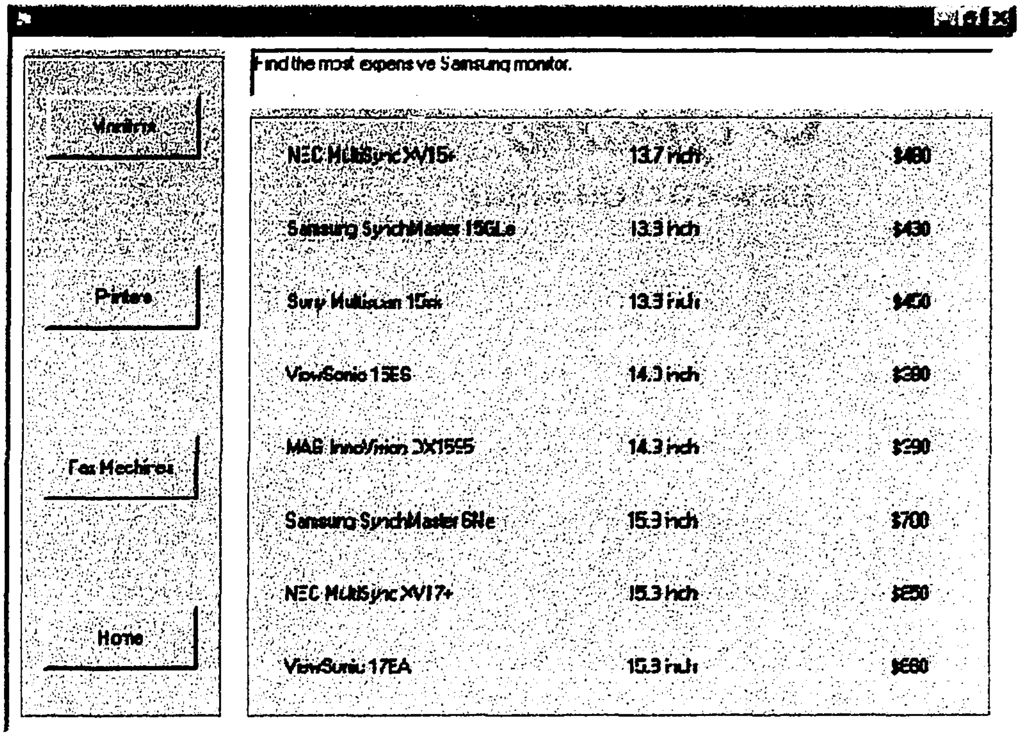

Find the fax machine with the largest cost per page.

Figure 1. Sample starting screen for both navigation aid conditions.

absent conditions is manifested in the interface. The navigation aid provided shortcuts to other areas of the software database program. Once a subject entered a particular area of the database by clicking on the corresponding button the difference between the navigation aid present and navigation aid absent conditions is manifested in the interface. An example of a subordinate screen in the navigation aid present condition is provided in Figure 2.

In the navigation aid *absent* condition, the three subject-specific buttons--Monitors button, Printers button, and Fax Machines button--are not visible except at the starting, or Home position, of the database. This starting screen is illustrated in Figure 1. An



Find the most expensive Samsung monitor.

| | | |
|---|---|---|
| NEC MultiSync XV15+ | 13.7 inch | $480 |
| Samsung SyncMaster 15GLe | 13.3 inch | $430 |
| Sony Multiscan 15sx | 13.3 inch | $450 |
| ViewSonic 15ES | 14.0 inch | $380 |
| MAG InnoVision DX1595 | 14.3 inch | $290 |
| Samsung SyncMaster 6Ne | 15.3 inch | $700 |
| NEC MultiSync XV17+ | 15.3 inch | $650 |
| ViewSonic 17EA | 15.3 inch | $660 |

Figure 2. Subordinate screen for the navigation aid present condition.

overview of the screen hierarchy is provided in Appendix A.

In the second main condition, Information Density, the amount of information density in the product area of the display was varied on a trial-by-trial basis. In the process of finding a target product for a particular trial, the subject had to extract information from the database that was displayed in high, medium or low information-density format. The occurrence of high, medium and low information density trials was counterbalanced in the testing procedure.

Tullis (1983) differentiates between overall and local density. Overall density concerns the total amount of free space available in an interface. Local density refers to the amount of space surrounding particular elements. The two are correlated. In the current research local density is defined and manipulated using the Tullis (1983) metric. Tullis (1983) suggests an index for local density to be an average percentage of characters in 88 spaces centered on the point of fixation given a standard definition of character spacing. This is based on five degrees of visual angle. To account for the differentiation of visual acuity with area within the five-degree diameter area around a character or point of fixation, Tullis used a linear weighting scheme to assign weights to the characters. The index was viewed as the average percentage of other characters near each character, with those closer being weighted more heavily. Although the software interface used in the current research is graphical, versus alphanumeric, the areas in which information density was manipulated are alphanumeric. Therefore Tullis' index of local density was used to define local density.

A five-degree circle is consistent with Danchak's (1976) choice of a 0.088-rad (5-deg.) circle as the maximum length of a displayed record. In the current research, five degrees of visual angle with a viewing distance of 24 inches translated into a 2.09-inch diameter circle. Tullis' linear weighting scheme was used to calculate local density for the three levels of information density used in the current study. For the low information density condition local density equaled 14%. For the medium and high information density conditions, local density equaled 36% and 68% respectively.

In the third condition, task complexity, a cognitive variable, the difficulty of the task was manipulated by modifying the amount of information used to specify task. In the low complexity condition, subjects were asked to find an item based on two criteria: peripheral category (Monitor, Printer, Fax Machine) and one of the two peripheral dimensions. In the medium complexity condition, subjects were asked to find an item based on three criteria: peripheral category, peripheral brand and one of the two peripheral dimensions. In the high complexity condition, subjects were asked to find an item based on four criteria: peripheral category, peripheral brand, item descriptor, and one of the two peripheral dimensions. The number of peripheral brands, item descriptors, and peripheral dimensions were all balanced within each of the three peripheral categories.

The overt response set contains one variable, i.e. mouse button clicks, for all trials and all conditions. The covert response set for all trials and for all conditions is eye movement. Subjects make overt responses by using their mouse to navigate the graphical software interface. Subjects make covert responses by moving their eyes while navigating the software interface displayed on the workstation's visual display terminal. Overt responses are recorded 30 times each second using a keystroke data logger (described below). Eye movement data is recorded 30 times each second using a corneal-reflection oculometer with head-mounted optics (described below). The sequence of presentation of the 18 within-subject conditions followed a counterbalancing schedule.

In the fourth and final condition, experience, subjects were randomly chosen then tested and divided into two groups. Group one consisted of subjects with less than six months of experience using a personal computer. Group two had more than three years of experience using a personal computer. The Computer Experience Questionnaire can be found in Appendix D. Five subjects were found to fit the criterion for group one and seven were found that fit the criterion for group three.

## Subjects

Subjects were 12 undergraduate students (6 male and 6 female). All subjects experienced all levels of the three within-subject independent variables. The two criteria affecting the selection of subjects from this pool was a requirement of 20/20 corrected vision and the amount of experience using software with a graphical user interface. Each subject was compensated with $20 for participating in the research. The American Psychological Association ethical principles nine and ten governing human subjects were observed.

### Apparatus

The software interface was created using Visual Basic in a Windows-95 environment. The software was displayed on a 21-inch color monitor with 16-bit color and a 640 x 480 pixel resolution driven by a Hewlett-Packard 66 megahertz Pentium computer workstation. A Microsoft mouse was used for navigation. Viewing distance was approximately 24 inches. Responses were made, as described above, by moving the mouse and depressing the left mouse button.

Subject lookpoint was measured using an ESP-ET-RH Remote/Head Mounted Eye Tracking System produced by ISCAN incorporated. The oculometer uses the corneal reflex technique to determine subject lookpoint (Young & Sheena, 1975). The system includes an ISCAN RK-426ESP Corneal Reflection Eye Tracker PC card. This card tracks the movement of the subject's eye within an image generated by the eye imaging subsystem, an ISCAN RK 520ESP Calibrator PC card. The Calibrator card calibrates the subject to a video scene and generates video overlay calibration points, an ISCAN Head Mounted Eye Imaging System with Head Tracking Sensor which consists of eye an imaging sensor, optics, an infrared illumination source and adjustable mechanical mounting. The ESP-ET-RH also included a Polhemus InsideTRAK magnetic position sensing electronics. These electronics, mounted on the subject's head, sent signals to the oculometer which allowed for measurement of head position with respect to a fixed magnetic source placed behind the subject, a Dell Pentium PC, Line of Sight & Target Intersection Software, three video monitors, and a VGA scan converter.

Subjects interacted with software on the stimulus computer. The stimulus computer recorded data on both the subject's lookpoint and the subject's mouse inputs to the system. The stimulus PC received eight bytes of data from the oculometer 30 times a second. The eight bytes included X and Y lookpoint coordinates on the stimulus plane (the video monitor of stimulus PC), and the plane number, which indicates for any given data string whether the oculometer was in or out of track.

The keystroke logging software ran in the background of the stimulus generating application, Visual Basic, and recorded 30 times a second the location of the mouse cursor and the state of the left mouse button. The logging software recorded the mouse data in synchrony with the lookpoint data being received via a serial cable link from the oculometer. The recorded data was written to a RAM drive until the end of the session when it was transferred to a hard drive for safer storage. The data logging software, created in Visual Basic, makes use of the 'DWSHK32.OCX' custom Visual Basic control from Desaware Software. The DWSHK32.OCX control provides access to Windows hooks to detect mouse clicks, again on a system-wide basis, before they are processed by the task. Both the Windows API and the OCX control events are triggered by the On_Comm event. This event is part of the MSComm communications control in Visual Basic. The event is fired every time the serial port on the stimulus PC receives data from the oculometer.

Time was measured on the host workstation using a timer function (TimerCount). The function is called from a dynamic link library called "toolhelp.dll" which can be found in the Windows/System directory in Windows 95. The timer was used in enhanced mode. TimerCount in enhanced mode uses the Virtual Timer Device to provide time stamps that are accurate to one millisecond.

### Task

Following a practice session, in which subjects used a database similar to the experimental database, each subject completed one full session that consisted of 36 trials. The navigation aid conditions were counterbalanced in their order of presentation to the subjects. The trial time was approximately 10 minutes.

### Task Instructions

Subjects were given on-line instructions for retrieving information from a hypertext-based information database containing information about computer peripherals. The instructions were written in recursive form; e.g., 'Find the least expensive fax machine'. There were 36 items for each subject to find. There were 6 randomized orders for the queries assigned to each of the twelve subjects. Each task description was presented serially where only one description was visible at a time. Each task description remained visible until the subject successfully completed the trial. A subject could not progress to a new trial until the subject completed successfully the previous trial. Task descriptions were presented at the top of the database interface. The description of the current task was always visible for the subject to refer to during the task.

When the subject acknowledged the confirmation screen by clicking the "Continue" button this made the confirmation screen disappear. The event also marked the beginning of the next trial. A description of the next item to find appeared in the task description space. A list of the task descriptions can be found in Appendix B. The contents of the product database are available in Appendix C.

### Computer Experience Questionnaire

This questionnaire was used to screen subjects for computer experience. This questionnaire is available in Appendix D.

### Procedure

Each of the 12 subjects served one time in each of the six independent conditions. The experiment was conducted in a windowless and sound attenuated room. Subjects were first read a set of identical formalized instructions. Verbatim text of the instructions is presented in Appendix E. The instructions described the nature the task they were being asked to perform and informed them of the dependent variables without revealing the hypotheses of the study. Subjects were given a short period to familiarize themselves with the software and the nature of the task they would be asked to complete (time >= 5 minutes).

The subject was calibrated before each session in order to measure accurately the subject's eye angular movement and lookpoint from the raw eye movement data collected by the oculometer. In order to reduce the amount of error head movement could have contributed to the oculometric data, subjects used a chin rest during the calibration and test procedure. The calibration involved three steps and lasted approximately five minutes. The first step is to achieve a good eye image using the RK-426OPC eye tracker. A good image of the eye must be obtained before proceeding to the remaining two steps of the calibration procedure. The experimenter uses subjective judgment viewing the video display of the eye image to adjust the head-mounted hardware and the software and obtain a good eye image. For the last two steps of the calibration procedure, the subject kept his head relatively still while he moves his eyes in response to two sets of five-point calibration patterns. The points are presented sequentially and the subject is instructed to look at each calibration point after which the experimenter registers the eye position. After looking at the specified points, the oculometer's calibration system computes a mathematical model that translates subsequent eye movement data into lookpoint data thirty times per second.

Following the calibration, the subject was asked to complete the work as quickly and efficiently as possible. The subject was then left alone in the room to complete the experimental procedure. Subjects responded with the preferred hand (right or left). Following the trial each subject underwent a short debriefing and received their stipend.

# RESULTS

## Keystroke Data Analysis

One type of keystroke data, left mouse button clicks, was recorded. Keystroke data was used to measure overt responses and define trials times in the task.

## Trial Time

Trial time is the time between the introduction of a query in the query window and a mouse click on the specified item in the database. There were 36 trial times measured with each subject. Trial times (TT) were analyzed using a General Linear Model Analysis of Variance procedure. The trial time (TT) data were analyzed with a 2 x 3 x 3 x 2 (Navigation Aid [present or absent] x Information Density [low, medium or high] x Task Complexity [low, medium or high] x Experience [low or high]) design. Experience was treated as a between-subject variable. The other three independent variables are within-subject variables.

Using an overall alpha level of .05 in the TT analysis, one main effect (Task Complexity), two two-way interactions (Information Density x Navigation Aid & Information Density x Task Complexity) and one three way interaction (Information Density x Task Complexity x Navigation Aid) achieved statistical significance. The results of the TT Analysis of Variance are summarized in Appendix A.

Task Complexity. A main effect of Task Complexity (low, medium or high) was found. Figure 3 illustrates this main effect. A Scheffe post-hoc test revealed that the mean TT in the low-complexity condition was significantly shorter than the mean TT in high-complexity condition. The mean TT in the medium-complexity condition was also significantly shorter that the mean TT in the high-complexity condition. The mean TT in

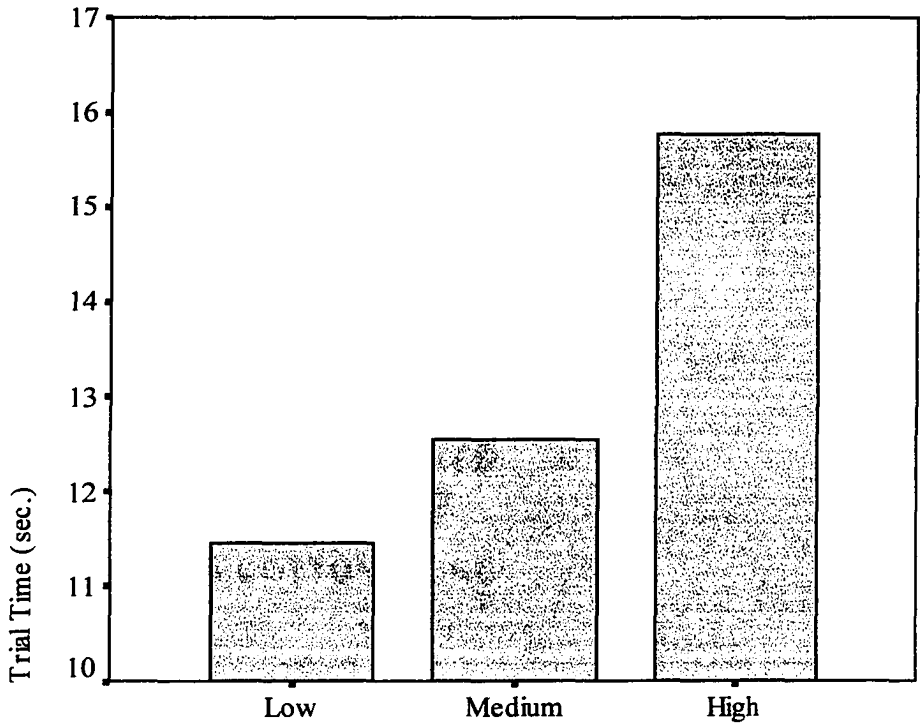


Figure 3. Task Complexity (TT)

the low-complexity condition was not significantly different than the mean TT in the medium-complexity condition.

Information Density x Navigation Aid. A significant two-way interaction (Information Density x Navigation Aid) was obtained. This interaction is illustrated in Figure 4. This interaction clearly shows that information density had no effect on trial time when the navigation aid was not present. In the trials where the navigation aid was present, the low and high information density conditions resulted in longer task times than the high and low information density trials in which the navigation aid was absent.

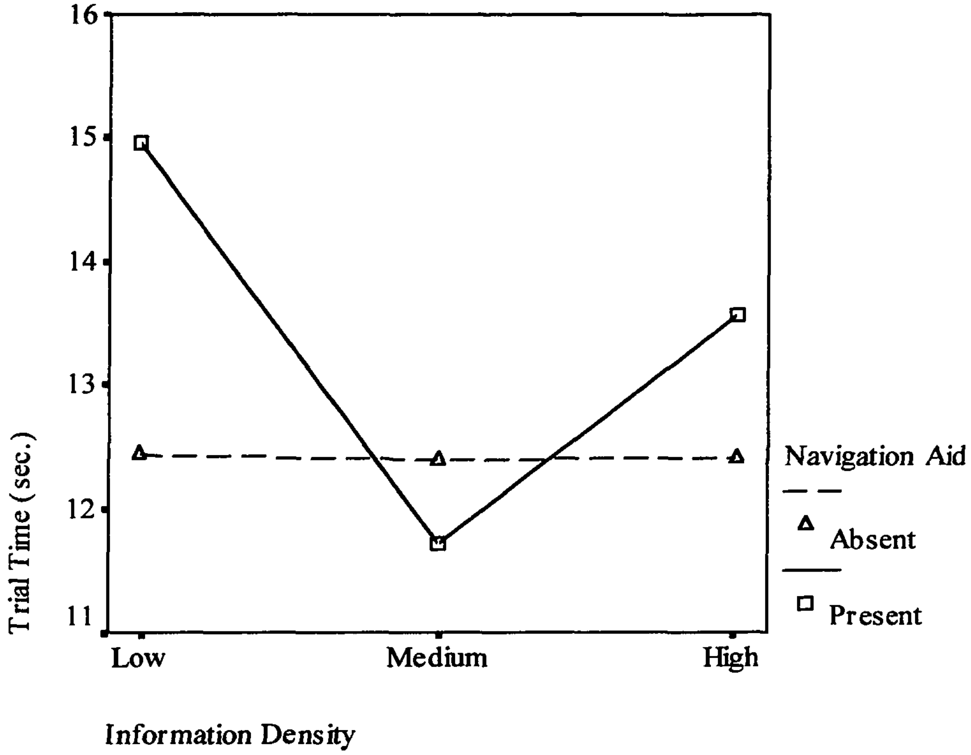


Figure 4. Information Density x Navigation Aid (TT)

An examination of the simple effects of the navigation aid within the information conditions showed that the navigation aid had a significant effect when information density was low, F(2, 20) = 18.376, p<.01 (Winer, Brown & Michaels, 1991). When information density was low, the presence of the navigation aid increased task time significantly. A further examination of the simple effects of the navigation aid within the information conditions showed that the navigation aid had a significant effect when information density was high, F(2, 20) = 3.857, p<.05. When information density was high, the presence of the navigation aid, again, increased task time significantly. No differences were present for the medium density conditions.

Information Density x Task Complexity. A significant two-way interaction (Information Density x Task Complexity) was obtained. This interaction is illustrated in Figure 5. This interaction shows that the effect of information density on trial time was different for medium complexity trials than for low and high complexity trials. While low information density inhibited task performance when the task complexity was low or high, it improved performance when the task complexity was medium.

An examination of the simple effects of information density within the task complexity conditions showed that the information density had a significant effect for low density trials, F(4, 40) = 3.301, $p<.05$. A Scheffe post-hoc test, however, revealed no significant difference in TT due to information density for low-complexity trials.

Simple effects analysis showed that information density had a significant effect for high density trials, F(4, 40) = 10.594, $p<.01$. A Scheffe post-hoc test indicated that the mean TT in the low-density condition was significantly longer than the mean TT in the medium and high-density trials.

Information Density x Task Complexity x Navigation Aid. The three-way interaction (Information Density x Task Complexity x Navigation Aid) yielded a significant result. This interaction is illustrated in Figure 6. This is a most interesting effect and this interaction shows that while task complexity had a significant effect for high density trials when the navigation aid was absent, task complexity had no effect on trial time for high density trials when the navigation aid was present.

A breakdown of this three-way interaction into its simple effects revealed that for trials in which the navigation aid was absent, task complexity had a significant effect,

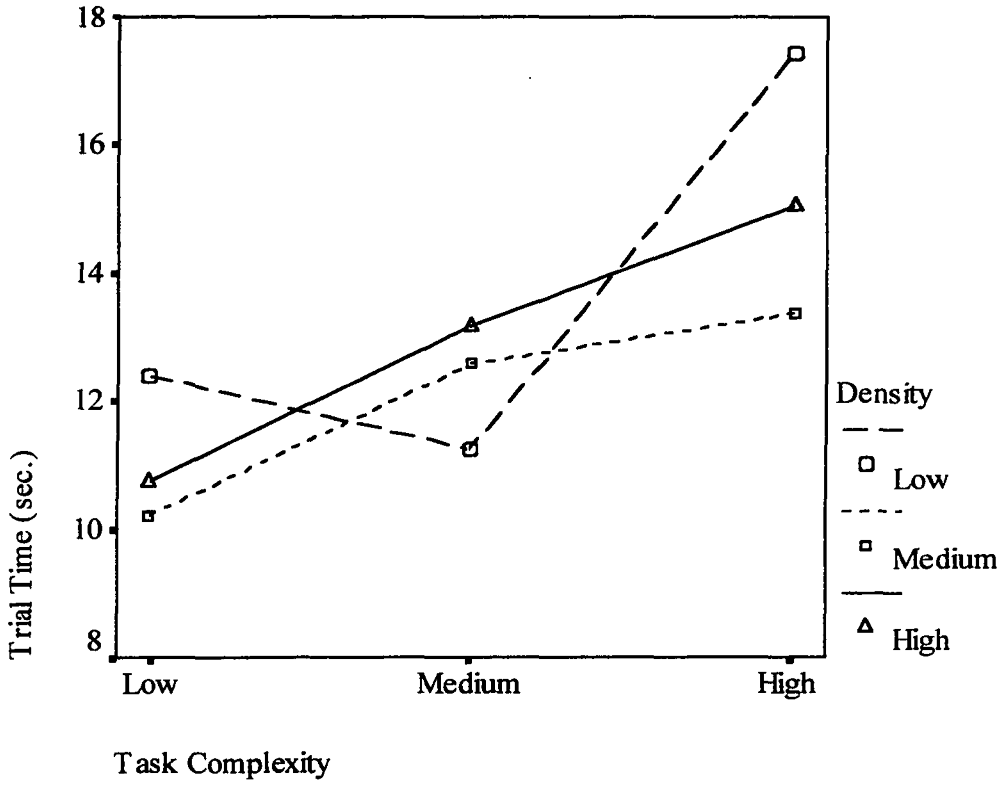


Figure 5. Information Density x Task Complexity (TT)

F(4, 40) = 31.101, p<.01 and the interaction between information density and task complexity had a significant effect, F(4, 40) = 119.075, p<.01. A Scheffe post-hoc test showed that for the trials in which the navigation aid was absent the mean TT for high complexity trials was significantly higher than the mean TT in both low and medium complexity trials.

A closer look at the information density x task complexity interaction for trials in which the navigation aid was absent revealed that task complexity had a significant effect for low density trials, F(4, 40) = 6.136, p<.01, and high density trials F(4, 40) = 24.165,

$p<.01$ but had no effect for medium density trials. A Scheffe post-hoc showed that the mean TT for high task complexity trials in the low information density environment was significantly greater than low and medium task complexity trials. The mean TT for high complexity trials in the high information density environment was again significantly higher than the mean TT in the low and medium complexity trials. The mean TT for medium complexity trials in the high information density environment was significantly higher than the mean TT in the low complexity trials

A further breakdown of this three-way interaction into its simple effects revealed that for trials in which the navigation aid was present, task complexity had a significant effect, $F(4, 40) = 9.222$, $p<.01$, information density had a significant effect, $F(4, 40) = 6.464$, $p<.01$, and the interaction between information and task complexity had a significant effect, $F(4, 40) = 4.884$, $p<.01$. A Scheffe post-hoc test showed that for the trials in which the navigation aid was present the mean TT for low information density trials was significantly higher than the mean TT for medium information density trials. In the trials where the navigation aid was present, the mean TT for high-density trials was significantly higher than the meant TT in for medium density trials. Another Scheffe post-hoc test showed that for the trials in which the navigation aid was present the mean TT for high density trials was significantly greater than the mean TT in both medium and low information density trials.

A closer look at the Information Density x Task Complexity interaction for trials in which the navigation aid was present showed that complexity had a significant effect for low density trials, $F(4, 40) = 8.954$, $p<.01$. A Scheffe post-hoc showed that the meant TT for high complexity trials was significantly greater than medium and low trials.

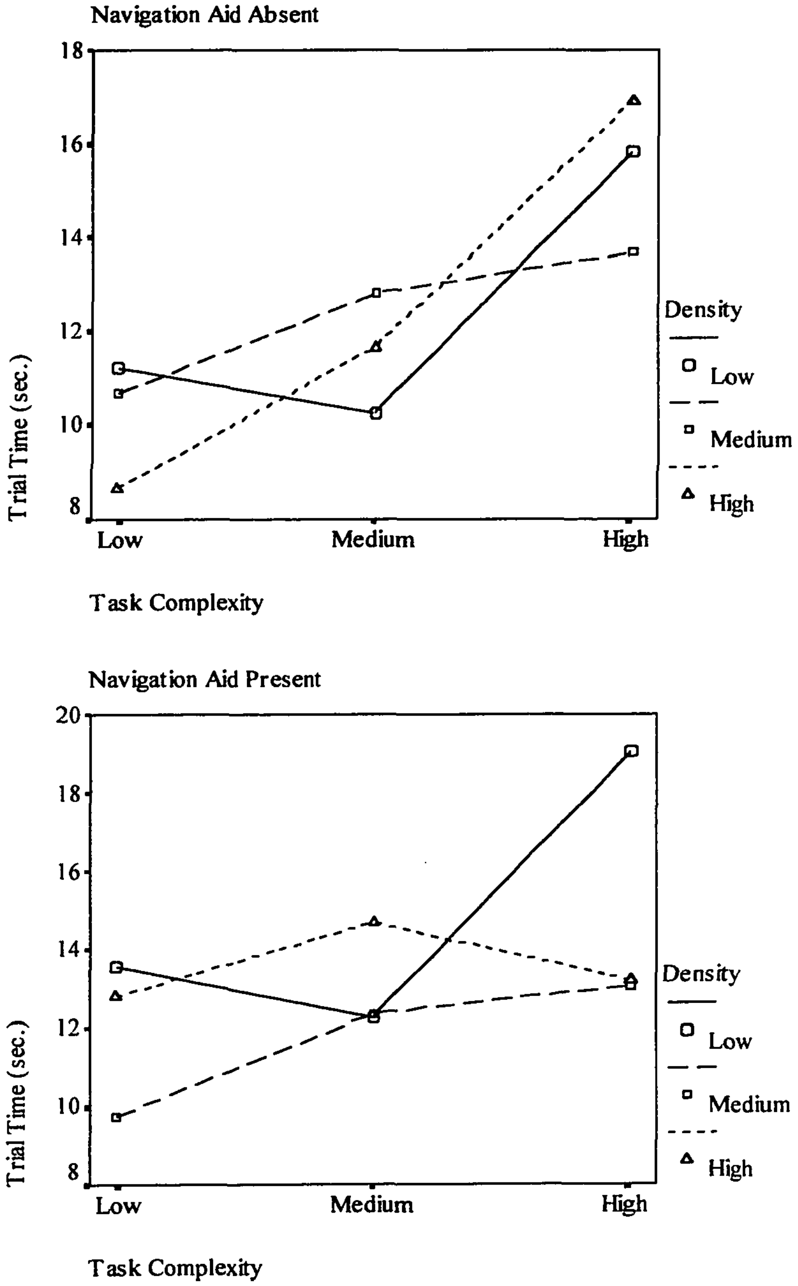


Figure 6. Information Density x Task Complexity x Navigation Aid (TT)

## Error Rate

Another dependent variable, error rate, was analyzed with an identical 2 x 3 x 3 x 2 Analysis of Variance design. Errors are defined as the total number of overt responses (mouse clicks) beyond the required number of mouse clicks necessary to complete a trial successfully. Error rate is the ratio of errors to trials for a particular condition. Using an overall alpha level of .05 in the TT analysis, one main effect (Task Complexity), one two-way interaction (Task Complexity x Experience) and one four-way interaction (Task Complexity x Information Density x Navigation x Aid Experience) achieved statistical significance. The results of the Error Analysis of Variance are summarized in Appendix B.

Task Complexity. A main effect of task complexity (low, medium or high) was found. Figure 7 illustrates this main effect. A Scheffe post-hoc test showed that the mean error rate in the low-complexity condition was significantly less than the mean ER in high-density condition. The mean ER in the medium-density condition was also significantly less that the mean ER in the high-density condition. The mean ER in the low-complexity condition was not significantly different than the mean TT in the medium-complexity condition.

Task Complexity x Experience. A significant two-way interaction (Task Complexity x Experience) was obtained. This interaction is illustrated in Figure 8. This interaction shows that while task complexity had no influence on the error rate of subjects with high experience, the ER jumped significantly for subjects with low experience when the task complexity was increased from medium to high.

An examination of the simple effects of task complexity within the experience levels showed that the task complexity had a significant effect for subjects with low

Figure 7. Task Complexity (ER)

experience, F(2, 20) = 24.292, p<.01. A Scheffe post-hoc test revealed that the ER for high complexity trials was significantly higher than the ER for both low and medium complexity trials. No other effects were significant.

Task Complexity x Information Density x Navigation Aid x Experience. A significant four-way interaction (Task Complexity x Information x Experience Density x Navigation Aid x Experience) was obtained. This interaction is illustrated in Figures 9 and 10. The most meaningful interaction the graph reveals is the influence task complexity has on error rate for subjects with high experience and high information density. Introducing the navigation aid in this situation resulted in an opposite effect on ER when

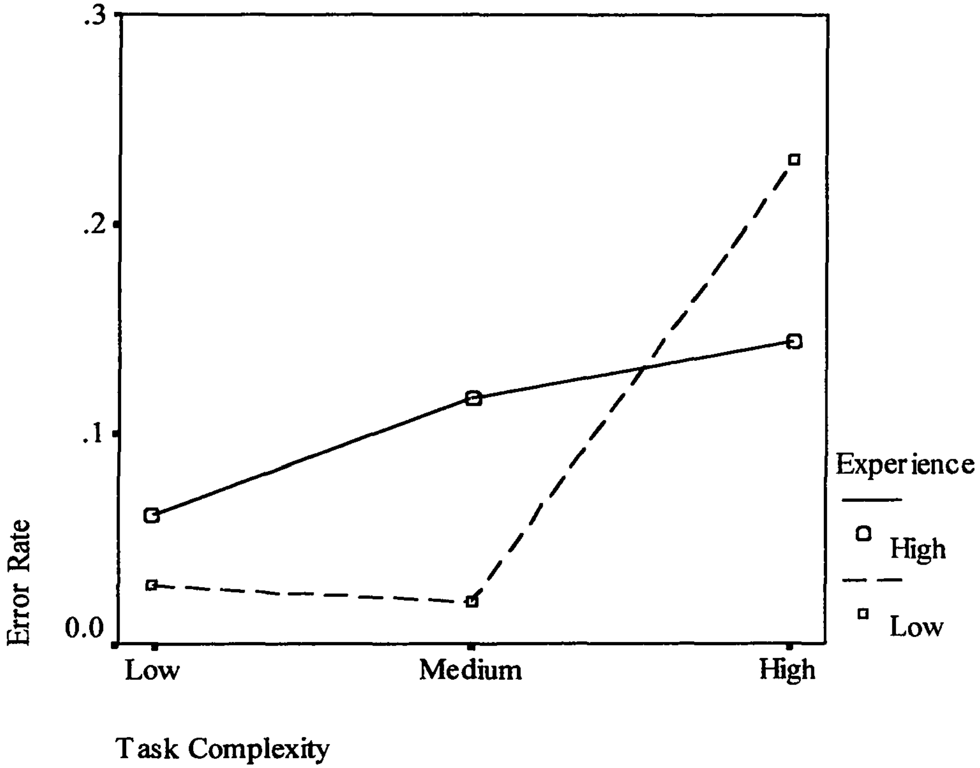


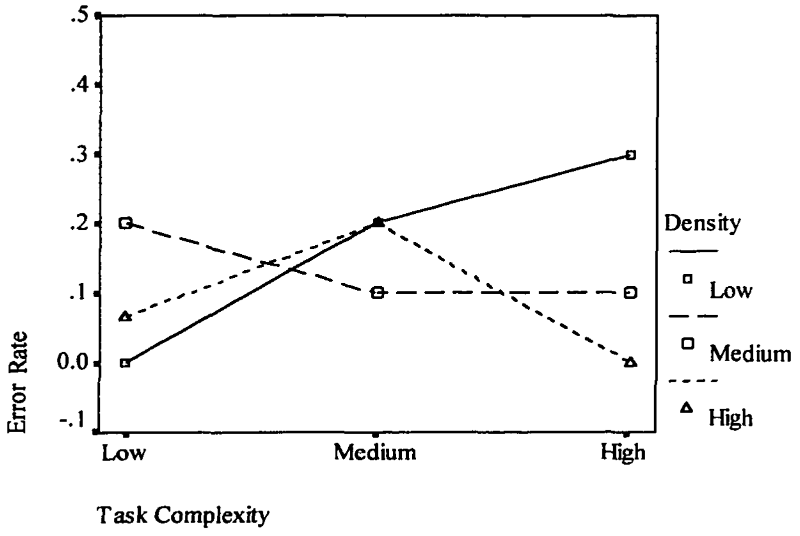
Figure 8. Task Complexity x Experience (ER)

the task complexity was increased from medium to high. While the ER increased significantly when the navigation aid was absent and the task complexity was increased, the ER went from relatively high to zero when the navigation aid was present and the task complexity was increased.

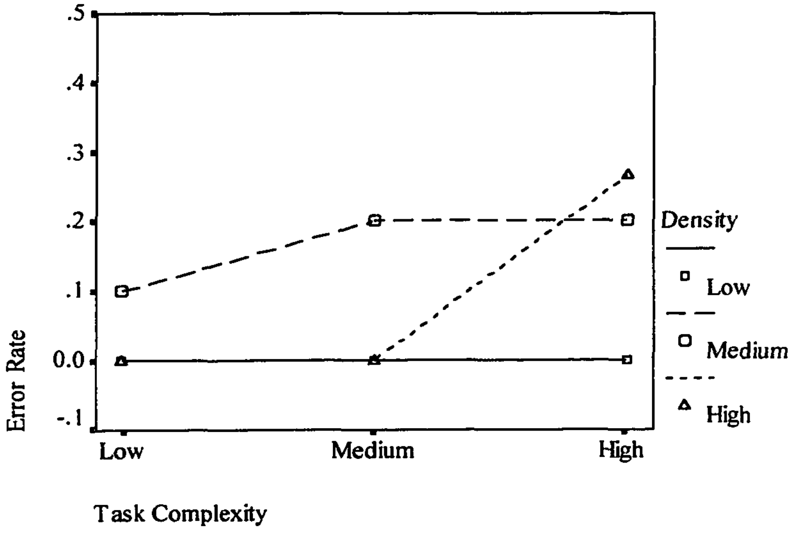
An examination of the simple effects of the Task Complexity x Information Density x Navigation Aid three-way interaction within the experience levels revealed that for subjects with low experience the effect of task complexity was significant $F(4, 40) = 22.278$, $p<.01$. A Scheffe post-hoc test showed that high complexity trials yielded a significantly higher ER than both low and medium complexity trials.

The simple effect of information density was also significant F(4, 40) = 4.376, p<.01. A Scheffe post-hoc showed that the ER for high-density trials was significantly higher than the ER for trials with medium information density. The simple effect of navigation aid was also significant for subjects with low experience, F(4, 40) = 3.969, p<.01. Low experience subjects yielded a significantly higher ER in trials where the navigation aid was absent. No other effects in the Task Complexity x Information Density x Navigation interaction were significant.

An examination of the simple effects within the Task Complexity x Information Density x Navigation Aid three-way interaction for subjects with high experience showed the effect of task complexity was significant F(4, 40) = 5.458, p<.01. A Scheffe post-hoc test showed that high complexity trials yielded a significantly higher ER than low complexity trials. The simple effect of navigation aid was also significant for subjects with high experience, F(4, 40) = 4.446, p<.01. High experience subjects had a significantly lower ER in absent navigation aid trials. No other effects were significant.

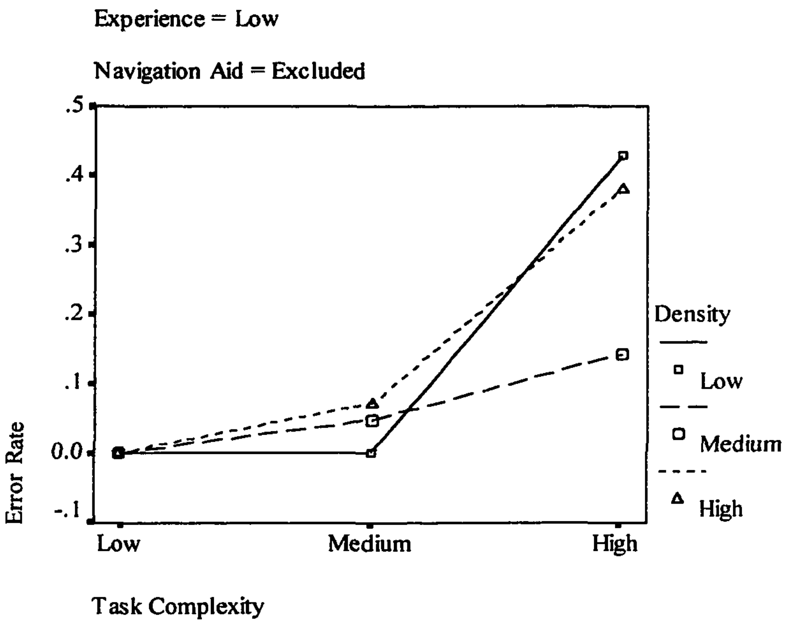

Experience = High

Navigation Aid = Included


Experience = High

Navigation Aid = Excluded

Figure 9. Task Complexity x Information Density x Experience x Navigation Aid (ER) (1 of 2)



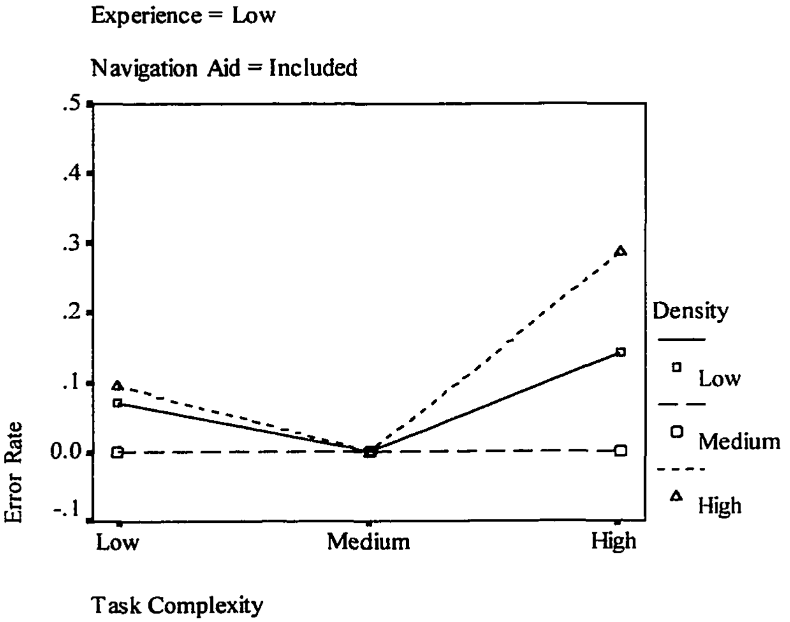


Figure 10. Task Complexity x Information Density x Experience x Navigation Aid (ER) (2 of 2)

Eye Movement Data Analysis

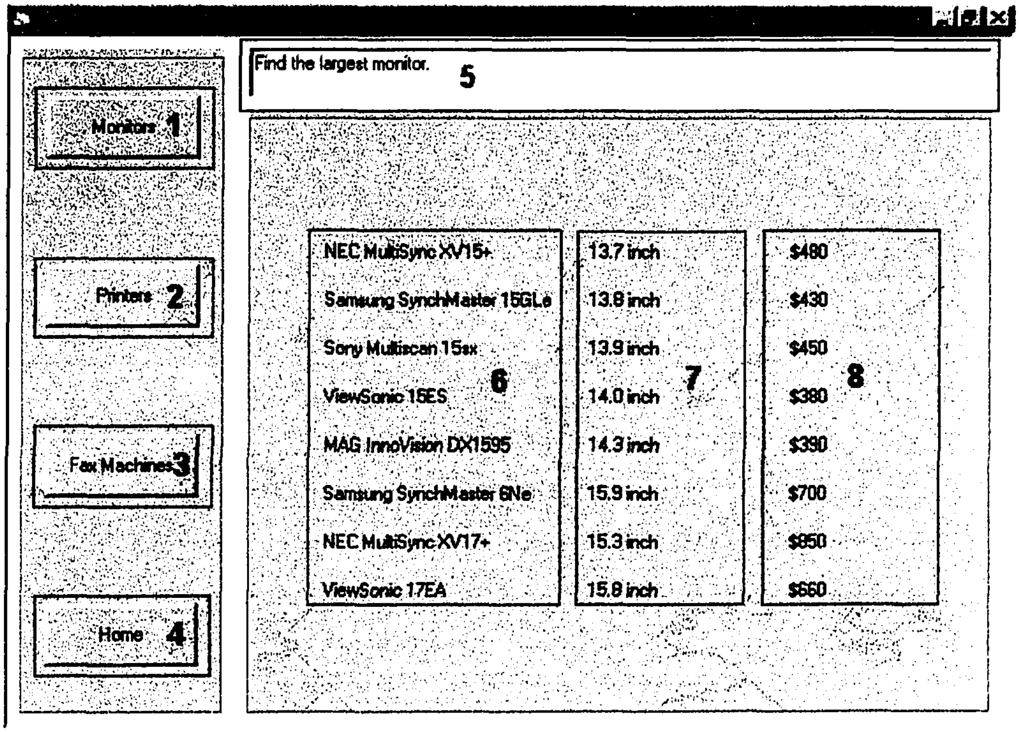
The eye movement data were analyzed using methods described by Harris, Glover & Spady (1986).

Originally scan-pattern contingency tables were going to be used to look at systematic differences in scan patterns as a result of changes in the independent variables. It was found through an early examination that manipulation of the independent variables resulted in only trivial differences in scan patterns and that the Chi-Square analyses revealed everything that the contingency tables would have revealed. Consequently the contingency tables and associated analyses have been omitted from these results.

Dwell Time

Dwell time is the amount of time spent looking at predefined areas of the display. Figure 11 illustrates the eight predefined areas of the display used in the dwell time analysis. Those display areas are: (1) Monitors, (2) Printers, (3) Fax Machines, (4) Home, (5) Query Box, (6) Product Identification, (7) Attribute A, (8) Attribute B.

One hundred forty-four dwell times were measured for each subject. Dwell times (DT) were analyzed using a General Linear Model Analysis of Variance procedure. The dwell time data were analyzed using an identical 3 x 2 x 3 x 2 (Information Density [low, medium or high] x Navigation Aid [present or absent] x Task Complexity [low, medium or high]) x Experience [low or high]) design. Using an overall alpha level of .05 in the DT analysis, one main effect (Task Complexity) and one two-way interaction (Information Density x Navigation Aid) achieved statistical significance. The results of the dwell time Analysis of Variance are summarized in Appendix C.



Find the largest monitor.

| | | |
|---|---|---|
| NEC MultiSync XV15+ | 13.7 inch | $480 |
| Samsung SynchMaster 15GLe | 13.8 inch | $430 |
| Sony Multiscan 15sx | 13.9 inch | $450 |
| ViewSonic 15ES | 14.0 inch | $380 |
| MAG InnoVision DX1595 | 14.3 inch | $390 |
| Samsung SynchMaster 6Ne | 15.9 inch | $700 |
| NEC MultiSync XV17+ | 15.3 inch | $850 |
| ViewSonic 17EA | 15.8 inch | $660 |

Figure 11. Eight display areas used in the dwell time analysis

Task Complexity. A main effect of task complexity (low, medium or high) was found. Figure 12 illustrates this main effect. A Scheffe post-hoc test showed that the mean dwell time in the high-complexity condition was significantly longer than the mean DT in both the low and medium-complexity conditions.

Information Density x Navigation Aid. A significant two-way interaction (Information Density x Navigation Aid) was obtained. This interaction is illustrated in Figure 13. This interaction shows that while information density had an effect on DT that resulted in a U-shaped function for trials in which the navigation aid was present, information density had no effect on trials in which the navigation aid was absent.

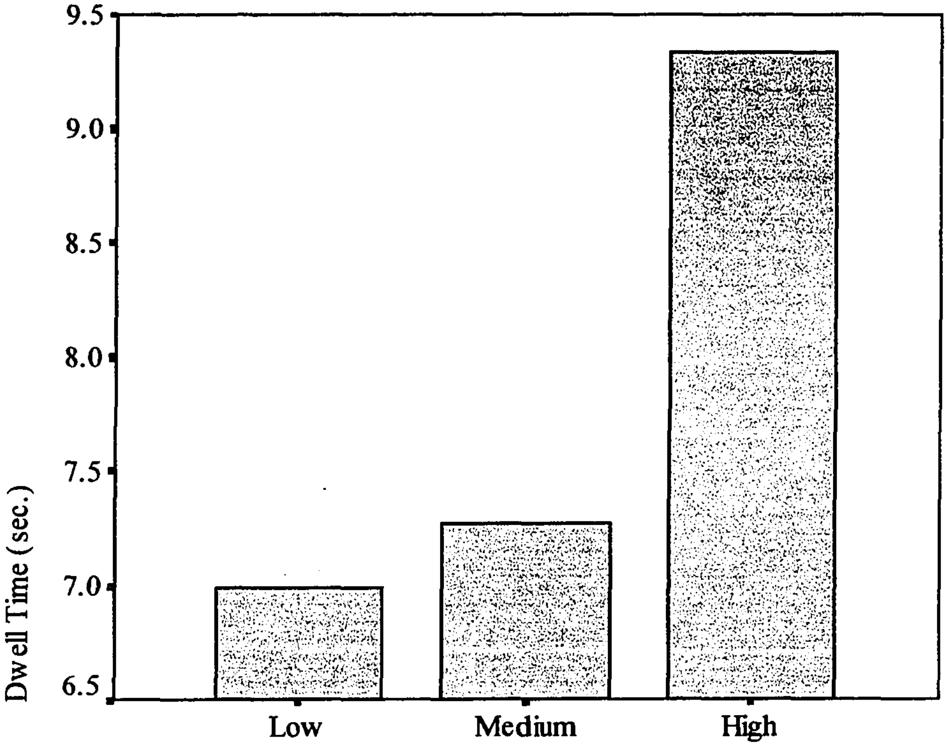


Figure 12. Task Complexity (DT)

An examination of the simple effects of information density within the navigation aid levels showed that the density had a significant effect for trials in which the navigation aid was present, F(2, 20) = 9.424, $p<.01$. The DT for trials with low information density was significantly greater than the dwell time for trials with medium information density. Also, the DT for trials with high information density was significantly greater than the dwell time in medium density trials. No other effects were significant.

## Dwell Frequency

Dwell frequency is defined as the number of times the eyes look at predefined areas of the display. Fitts et al. (1950) interpreted dwell frequency to be a reflection of the

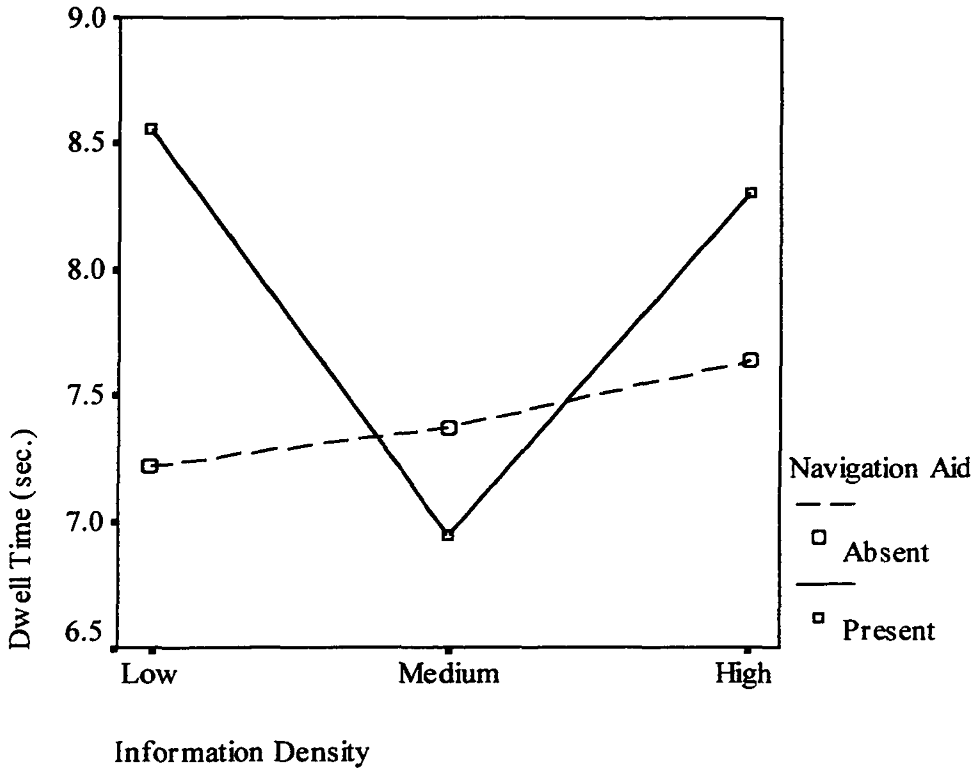
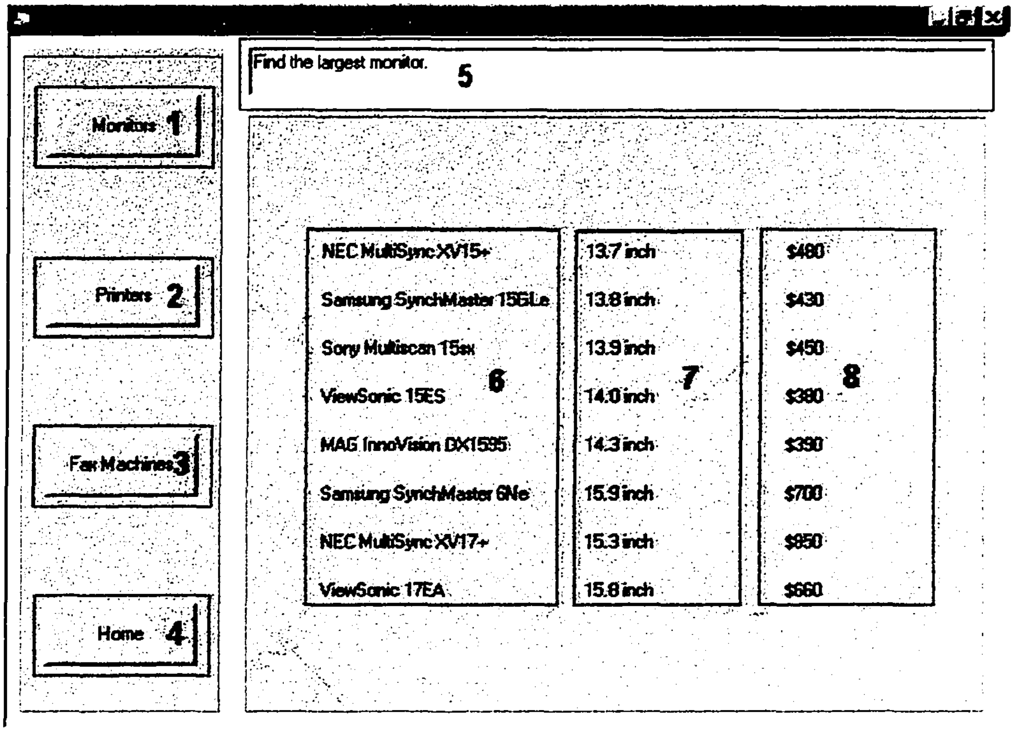


Figure 13. Information Density x Navigation Aid (DT)

importance of an object. Importance is a subjective construct.

Figure 14 illustrates the eight displays areas used in the analysis. Those display areas are: (1) Monitors, (2) Printers, (3) Fax Machines, (4) Home, (5) Query Box, (6) Product Identification, (7) Attribute A, (8) Attribute B. The number of times the subject looked at each display element during each of the 36 trials was measured.

A Chi-Square procedure was used to compare the effects of information density, experience, navigation aid and task complexity on dwell frequency. Four separate Chi-Square analyses were performed, each testing the effects of a different independent



Find the largest monitor.

| | | |
|---|---|---|
| NEC MultiSync XV15+ | 13.7 inch | $480 |
| Samsung SynchMaster 155Le | 13.8 inch | $430 |
| Sony Multiscan 15sx | 13.9 inch | $450 |
| ViewSonic 15ES | 14.0 inch | $380 |
| MAG InnoVision DX1595 | 14.3 inch | $390 |
| Samsung SynchMaster 6Ne | 15.9 inch | $700 |
| NEC MultiSync XV17+ | 15.3 inch | $850 |
| ViewSonic 17EA | 15.8 inch | $660 |

Figure 14. Eight display areas used in the dwell frequency analysis

variable. Three of the Chi-Square comparisons were found to be significant.

Information Density. A significant effect for information density level (low, medium or high) was found Chi-Square ($p < .05$, 14) = 32.025. This effect is illustrated in Figure 15.

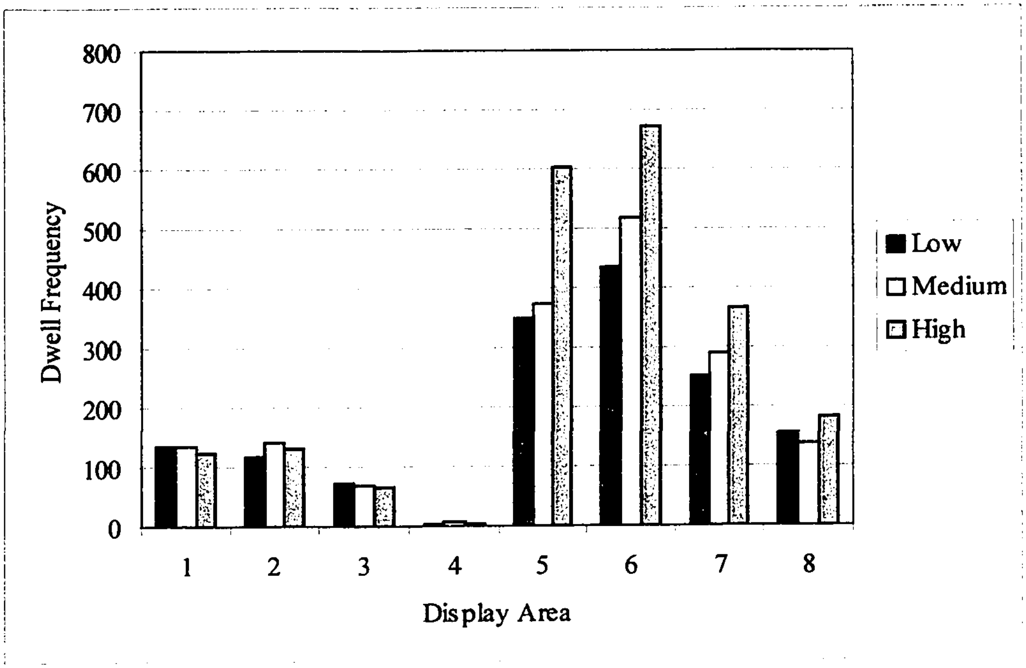


Figure 15. Information Density x Display Area (DF)

Experience. A significant effect for experience (low or high) was found Chi-Square ($p < .05$, 7) = 19.939. This effect is illustrated in Figure 16.

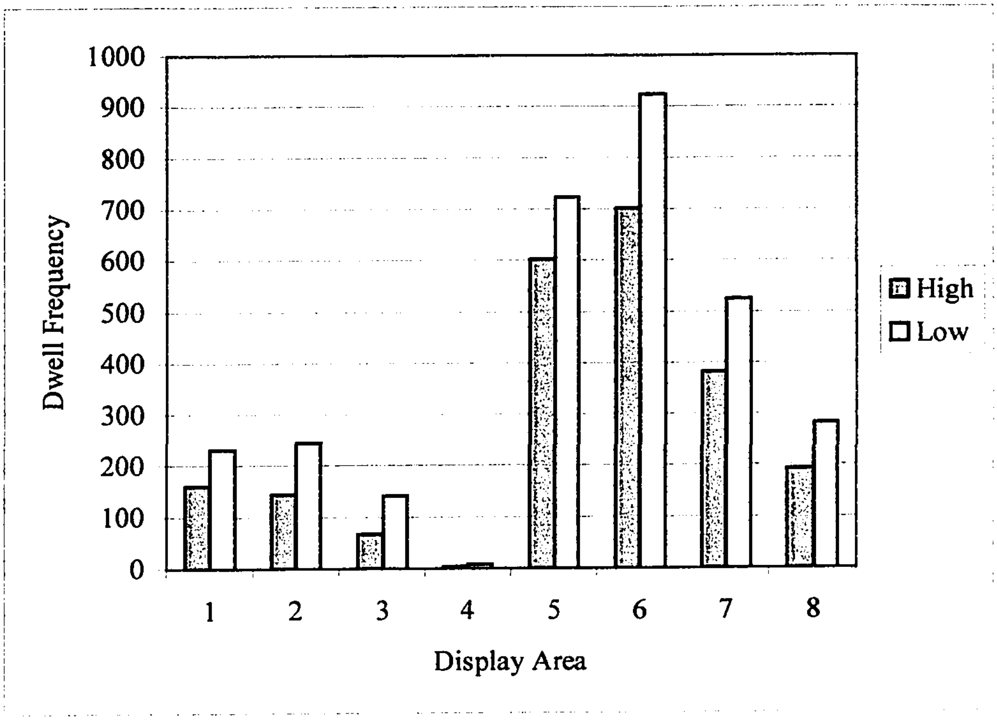


Figure 16. Experience x Display Area (DF)

Task Complexity. A significant effect for task complexity (low, medium or high) was found Chi-Square (p< .05, 14) = 55.934. This effect is illustrated in Figure 17.

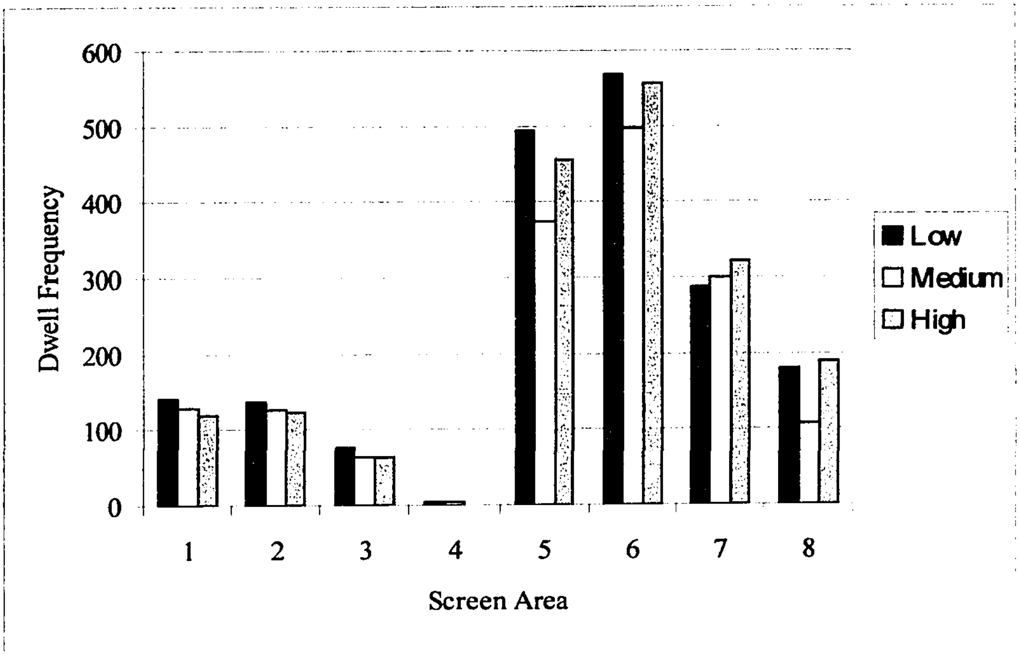


Figure 17. Task Complexity x Screen Area (DF)

# FINDINGS AND INTERPRETATIONS

A basic assumption driving this research was that if one can quantify and accurately describe scanning behavior and how it relates to performance then one can design better information displays. Tannas (1985) emphasizes, "It is critical for the display designer and systems engineer to remember at all times that the ultimate purpose of any visual display technology is to provide useful and appropriate information to the person(s) using the display" (p. 57). The current research furthers our understanding of how computer operators use their eyes to extract information from visual displays, an often-overlooked aspect of human-computer interaction. It has also increased our understanding of how ocular behavior relates to more conventional measures of performance. Lastly the current research furthers our understanding of the relative role cognitive characteristics and display characteristics have on both human performance and ocular behavior. This research was designed to test a number of hypotheses. Support for these hypotheses in light of the results of this research is considered next. A more general discussion of the results follows.

Hypothesis 'A' states that displays with high or low information density are more difficult to interpret and result in lower rates of information transfer than display areas with moderate information density. In order to test this hypothesis task time and dwell time were used as dependent measures.

The information density variable in the current research was based on a local density index defined by Tullis (1983). In the analysis of the task time data, information density did not have a main effect as expected. Information density did, however, interact with the navigation aid and complexity variables.

The interaction between information density and navigation aid shows that while information density had no effect while the navigation aid was absent, the display characteristic had an effect in those trials in which the navigation aid was present. These results for the trials in which the navigation aid was absent exhibit an inverted U-shape function similar to what Landis, Slivka & Jones (1967) proposed. Their proposed function was based on the theory that the relationship relating quality of performance and display density has an inverted-U shape. At low levels of information density, raising the density enhances performance while at high levels it inhibits performance. This implies that an optimal level exists. For this inverted U shaped function observed in this interaction the optimal level appears to be in the medium information density condition.

An obvious question is why didn't the function carry through to the trials in which the navigation aid was absent? Holahan, Culler and Wilcox (1978) found a positive relationship between the level of visual distraction in a display space and reaction time. They showed that the ability to locate and respond to a target sign in a cluttered display was directly inhibited by the proximity of other stimuli in the field of view. One explanation for this interaction between information density and the navigation aid is that, similar to Holahan, Culler and Wilcox's (1978) results, the presence of the navigation aid in the target screen inhibited performance in those trials where density deviated from the optimal level alluded to in Landis, Slivka & Jones's (1967) research.

Information density also interacted with task complexity. Interestingly, density had a significant effect for low and high complexity trials, but not medium complexity trials. Although the simple effect of information density for low complexity trials was significant there were no differences found between the task times associated with different levels. If the task complexity variable is thought of as a construct that focuses on the concept of difficulty then it can be thought of as a mental workload variable (Moray, 1979). The effect of information density in the high task complexity trials suggest that cognitive workload may have influenced performance. Although low density trials were significantly longer than medium or high density trials when the task complexity was high, the lack of difference between the medium and high density trials suggests that this interaction may be limited to low information density conditions.

Information density did not have a main effect in the analysis of the dwell time data. Information density did, however, interact with the navigation aid variable in a two-way interaction as it did in the task time data analysis. Again, as was found in the task time data, information density had an effect on dwell time in trials where the navigation aid was present but had no effect in trials where the navigation aid was absent. For those trials in which the navigation aid was present, the effects of the information density showed an inverted U-shaped function similar to the one observed in the Task time data (Landis, Slivka & Jones, 1967). It could be argued that the mere presence of the navigation aid buttons increased the overall information density. In a pure sense it would have increased the information density of each of the three levels by a constant amount meaning that the relative levels of information density remained constant but the absolute levels all increased. Perhaps the absolute increase in information density through the presence of the navigation aid was necessary for the cognitive workload variable to influence task time or dwell time. This hypothesis could help to explain why the two-way interaction between information density with navigation aid was found. Future research would be required to test this hypothesis.

Information density had a significant effect on the frequency of dwells. Dwell frequencies on the two areas with the highest frequency of dwells, the Product Identification data field and the Query window, showed an inverted U-shaped function (Landis, Slivka & Jones, 1967). In the two most visually accessed areas of the display, trials with low and high information density produced greater number of dwells than the medium information density trials. This function appeared in only one of the other six, less frequented, screen areas.

These results appear to support other research that shows information density has little effect on performance time (Wickens & Andre, 1990). High-density environments retard performance but also require less visual scanning, with more information captured per fixation. Lower display density results in greater scanning distances but less performance attenuating clutter. Thus the two factors, visual scanning and visual clutter, essentially trade of with one another as target dispersion changes. In the current research information density did not affect time-based measures of performance, it did influence ocular behavior. An optimal level of information density that allows the human operator to extract information with the least number of dwells may exist. If differences in information density do influence performance as indicated through time or error-based criteria, its role as a predictor of performance is limited.

Hypothesis 'B' stated that the arrangement of display elements in the visual field influences performance. Hypothesis 'C' stated that the arrangement of display elements will influence strategies employed by the subjects. Support for these two hypotheses in the results of the current research are considered in conjunction.

The navigation aid variable in the current research was based on the presence or absence of a navigation aid consisting of three buttons. In the analysis of the task time data, navigation aid did not have a main effect as expected. The status of the display element in the software did influence performance but this depended on the status of other variables in the design. This influence is seen in the Navigation Aid x Information Density interaction. As referred to above in the trials where the navigation aid was present, the information density variable had an effect on performance. But information density had no effect in those trials where the navigation aid was absent.

Reviewing other related research findings may facilitate interpretation of these results. Holahan, Culler and Wilcox (1978) found that the number and proximity of visual distractors in the visual field had significant effects on RT. Wickens and Andre (1990) found that when focussed attention required the close spatial proximity of distractors, the distracting elements disrupted performance. Eriksen (1995) showed that the number of dwells increased when the number of irrelevant stimuli in the visual field increased. Eriksen hypothesized that a relevant distracter at one time may become an irrelevant distracter at another time.

The only influence the navigation aid variable had in the dwell time data was in the information density by navigation aid two-way interaction. Here again as was seen earlier in the results of the Task time data, the trials where the navigation aid was present, information density had a significant effect. As was hypothesized above, the mere presence of the navigation aid could have increased each level of overall screen information density by a constant amount. The navigation bar may have been perceived at relevant at times and irrelevant, or a distracter, at other times lending credence to the notion that at some points it was a distracter impeding performance. More research would be needed to test this hypothesis.

Contrary to expectations, the navigation aid variable had no effect on dwell frequency. This could be interpreted to mean that the presence or absence of this display element had no effect on the strategy adopted by the subjects to complete the tasks. The low frequency of dwell counts for the buttons that constitute the navigation aid show that it was not fixated upon often relative to the other areas of the display. From an information theory perspective this suggests that the subjects did not perceive the navigation aid as a source that could reduce their task-induced uncertainty (Shannon, 1948).

Senders (1983) and Van Delft (1987) argue that visual sampling is independent of instrument arrangement. Donk (1994), like Fitts et. al (1950), considers spatial arrangement to be one of the major sources of variance in sampling behavior. The results of the current research appear to support the former view of Senders and Van Delft. There was no difference in the visual sampling frequency of the subjects due to the status of the navigation aid variable. But it cannot be simply stated that visual scanning behavior is dependent or independent of display characteristics, like many aspects of psychology, it depends on a variety of factors. Cognitive and display characteristics most likely play different roles depending on the environmental and cognitive characteristics of the situation. Attempts at modeling human monitoring behavior have been developed but in this information age, models that predict interactive performance need to be developed (Senders, 1983; Sheridan, 1970; Stein & Werwerinke, 1983).

The results relating to the navigation aid variable can be considered in the light of queuing theory, one of the models of monitoring performance (Senders, 1983). Queuing theory has been used in the analysis of systems with one or more service channels and some customers, or operator's attention, who serially service the channels. In this conceptual framework, an instrument, or display element, is serviced until the uncertainty about it is reduced to zero. At that time the customer (attention) leaves the service channel to engage the next service which has the highest probability of reducing uncertainty. In the context of the current research, the navigation aid was not serviced often by the customers (each subject's attention). This assumes that each subject's attention was highly correlated with the target of visual fixation. This contention is supported in other research (Fitts et. al, 1950; Norman , 1968; Posner, Snyder and Davidson, 1980). In the light of queuing theory the navigation aid was not serviced because it did not offer much help in reducing the uncertainty introduced with each task.

Hypothesis 'D' stated that the experience influences strategy. Experience, as it is defined in the current research, had no effect on either of the time-based measures: task time or dwell time. Experience did interact with task complexity to yield a significant two-way interaction in the error rate analysis. While task complexity had no effect on the error rate of subjects with high experience it had a significant effect on the error rate of subjects low experience subjects. The higher error rate for the subjects with low experience could be a symptom of a faulty strategy or mental model of the system. Although the subjects with low and high experience had spent the same amount of practice with the custom database application that was used in the experimental task, their differing levels of computer experience may have influenced the amount free cognitive resources each had available to apply towards the tasks. If, from either a single or multiple resource theory perspective, some of the perceptual-motor skills associated with basic computer interaction had been more automated for the subjects with high experience, those tasks would draw less resources then would be required for the low experience subjects (Boff, Kaufman & Thomas, 1987). This could help to explain the rise in error rate for the low experience subjects due to the change in task complexity, or cognitive workload, from medium to high.

An examination of the dwell frequency results indicates that experience had a clear and consistent effect on the dwell frequency data for each of the eight areas of the display. It should be noted that inferring strategy or lack of it, from dwell frequency patterns is a subjective interpretation. For each of the eight pre-defined areas of the display, the number of dwells required by the subjects with low experience was consistently greater than the number of dwells required by those with the high experience. Differences in dwell frequencies between these two groups of subjects engaged in the experimental task implies differences in attention which reflects differences in cognitive functioning. The greater number of fixations required could be indicative of a less efficient search strategy by the low experience subjects, or less information transferred per dwell. Since experience had no effect on task time the greater number of dwells by the less experienced subjects may have compensated for the lower amount of information transmission per dwell by increasing the greater number of dwells. This could help to explain the lack of effect of experience on the time-based measures.

This interpretation supports previous research that suggests eye movement parameters correlate with the information gathering strategies of the subjects (Antes, 1974; Mackworth & Morandi, 1967). This supports the research findings of Unema and Rotting (1985) who found that less experienced subjects had longer dwell times than more experienced subjects. Russo and Rosen (1975) who hypothesized that experts extract more information from a dwell but they also extract this information at higher rate so the dwell time for the expert and the novice is not different. Krappman (1995) found that successful subjects used a more selective information gathering strategy than unsuccessful subjects although it should be noted that Krappman divided his subject into the these groups post-hoc. Wickens (1992) suggests that dwell lengths are related to the difficulty of information extraction. Wickens also argues that the dwell length and the amount of information extracted are correlated but not perfectly.

Both Levy-Schoen (1981) and Wickens (1992) contend that scan patterns are dominated by cognitive factors and that display characteristics play a less significant larger role in determining scan patterns. Other research has demonstrated the role of cognitive factors underlying scan patterns and fixations (Boff, Kaufman & Thomas, 1987; Stark & Ellis, 1981). The larger number of dwells for less experienced subjects may indicate a less optimal search strategy. This interpretation should be considered in conjunction with the current research that demonstrated experience did not affect trial time or dwell time.

Wickens (1992) also argued that scan patterns are a reflection of a mental model. Differences between experts and novice fixation patterns indicate how the mental model or search strategy of the novice departs from that of the expert. If Wickens' argument is true and scanning behavior reflects the subject's mental model of the environment, then scanning behavior can also be thought of as an index of the subject's information needs. A less refined mental model may results in a less optimal search strategy. The greater number of dwells, reflected in the significant Chi-Square analysis of the dwell frequency data, suggests that the dwell frequencies of the less experienced subjects may have been a function of a database model that was less refined than the model used by the high experience subjects (Stark & Ellis, 1981).

In summary of hypothesis D, the results of the current research showed that although experience did not influence either of the time-based measures, experience had a significant effect on oculomotor behavior and influenced the error rate of the less experienced subjects. Although these results, and other research, appear to support the hypothesis that experience influences strategy, still further research is needed to understand better the relationship between experience and strategy (Jones, 1985).

Eye movements are a product of both environmental and internal, or cognitive, factors (Harris & Spady; 1985; Wickens 1992). While the relative amount of influence these variables have on ocular behavior is debated in the scientific literature, research has shown that areas in the visual field with high information content attract fixations (Tullis, 1983; Wickens, 1987). Because scan paths over same visual stimuli will vary according to changes in experience and goals, information transmission is therefore not a static property but varies in accordance with situational characteristics. It can be argued that much of visual search is internally driven by cognitive factors, as this research has shown. Out of the four independent variables controlled in the current research, task complexity, a cognitive variable clearly had the most powerful effect both the time-based measures of performance and the oculometric measures of performance. Task complexity yielded a main effect in the task time data, the error rate data, the dwell time data and the dwell frequency data where an increase in task complexity yielded increases in task time, error rate, dwell time and dwell frequency. Changes in mental workload had a greater effect than experience or the two time-base measures, information density or navigation aid.

Other research reports support the influential role of cognitive load on both performance and oculometric behavior (Donk, 1994; Senders, Elkind, Grignetti & Smallwood, 1964; Stark & Ellis, 1981; Yarbus, 1967). The current results do not support the position of object hypothesis advocates who argue that lower order aspects of physical stimuli in the environment generally determine eye movements (Didday & Arbib, 1972).

In the current research oculometric data offered insights beyond that which was available through conventional measures of performance. Combining, for example, task time data with dwell frequency data offered a richer account of how the independent variables under study influenced the operator. Take away either source of data, task time or dwell frequency and the depth of the account decreases. The oculometric data provided insights into how the operators distributed their attention and accomplished their tasks under the varying conditions while the conventional measures of performance provided standard measures to use when comparing the results to other research.

The relationship of attention and eye movements is an old question in psychology. While research has shown that shifts in attention occur independent of eye movements, the correlation between the two is very high (Eriksen & Hoffman, 1972; Jonides, 1983). These experimentally accessible quantities, argued by some to be controlled by cognitive models, provide a unique source of data inaccessible using other measures. There are weaknesses in interpreting some oculometric measures. For example, interpreting the meaning of a longer dwell can be difficult. A long dwell may reflect slower information transfer, more information being transferred, or staring (Harris et al., 1986). Norman (1968) used the metaphor of a spotlight to describe attention. Since eye movement is often highly correlated with this spotlight, tracking the scan patterns of subjects in the current research provided a unique look into the cognitive activity of the subjects.

Considering how the subjects may have developed mental models to complete the prescribed tasks may provide insight into the results of the current research. Moray (1990) argues that mental models are generally not accessible to consciousness but it is hypothesized they guide ocular behavior. Environment influences and shapes cognition that in turn guides the operator's interaction with the environment (Moray, 1990). The attention demand of a display is related in the current research to the probability that the display element will yield information. While a low bandwidth signal source or an area with low probability for yielding useful information may not attract much visual attention, it has a greater chance of influencing attention than if the element did not exist or was not visible at all. From an information design standpoint, Moray argues that displays do not just provide information but also control attention. The current research supports this argument to a degree but also underscores the strong influence of task characteristics and cognitive factors on aspects of performance.

The number and complexity of information displays is increasing in our information age. The display community is becoming increasingly aware of human interface problems that arise with the pervasiveness of display technology. The current research helps us to better understand how computer operators use their eyes to extract information from visual displays, how ocular behavior relates to more conventional measures of performance and the role cognitive characteristics and display characteristics have on human performance. The author's hope is that this research, in conjunction with future research, can be used to help develop theory based, and therefore generalizable, display design principles.

# REFERENCES

Abernathy, B. (1988). Visual search in sport and ergonomics: Its relationship to selective attention and performer expertise. Human Performance, 1(4), 205-235.

Antes, J. R. (1974). The time course of picture viewing. Journal of Experimental Psychology, 103, 62-70.

Boff, K., Kaufman, L. & Thomas, J. (Eds.). (1987). Handbook of Perception and Human Performance. New York: Wiley.

Bush, V., (1945, July). As we may think. The Atlantic Monthly, 176, 101-108. Reprinted and discussed in Interactions, 3(2), Mar 1996, 35-67.

Cavanagh, P. & Antis, S. (1991). The contribution of color to motion in normal and color-deficient observers. Vision Research, 31, 2109-2148.

Comstock, J. R., Jr. (1983). Oculometric indices of simulator and aircraft motion. (Doctoral dissertation, Old Dominion University, Norfolk, VA).

Danchak, M. M. (1976). CRT displays for power plants. Instrumentation Technology, 23(10), 29-36.

Demaio, J., Parkinson, S., Leshowitz, B., Crosby, J., & Thorpe, J. A. (1976). Visual scanning: Comparisons between student and instructor pilots (Report No. AFHR-TR-76-10, AD-A02363.4). Williams Airforce Base, AZ: Air Force Human Resources Laboratory.

Didday, R. L., & Arbib, M. A. (1972). Eye movements and visual perception: A 2 visual system model. (COINS Technical Report No. 73C-9) Amherst, MA: University of Massachusetts.

Donk, M. (1994). Human monitoring behavior in a multiple-instrument setting:

Independent sampling, sequential sampling or arrangement-dependent sampling. Acta Psychologica, 86, 31-55.

Drury, C. (1975). Inspection of sheet metal: Model and data. Human Factors, 17, 257-265.

Dumas, J. S. & Redish, J. C. (1994). A practical guide to usability testing. Norwood, NJ: Ablex.

Eriksen, C. W. (1995). The effect of extraneous stimuli on eye movements. In J.M. Findlay (Ed.), Eye Movement Research (pp. 315-336). Amsterdam: Elsevier Science.

Eriksen, C. W. & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. Perception and Psychophysics, 12, 201-204.

Faraday, P. & Sutcliffe, A. (1997). Designing effective multimedia presentations. CHI 97 Conference Proceedings, 272-278.

Fitts, P. M., Jones, R. E. and Milton, J.L (1950). Eye movements of aircraft pilots during instrument landing approaches. Aeronautical Engineering Review, 9, 1-5.

Flagg, B. N. (1978). Children and television: Effects of stimulus repetition on eye activity. In J.W. Senders, D.F. Fisher, & R.A. Monty (Eds.), Eye Movements and the Higher Psychological Functions. (pp. 279-291). Hillsdale, NJ: Lawrence Erlbaum.

Gale, A. G. & Worthington, B. S. (1984). Image viewing habits and the false diagnosis. Image Viewing, 6, 130-136.

Giroux, L. & Belleau, R. (1986). What's on the menu? The influence of menu content on the selection process. Behavior and Information Technology, 5, 196-172.

Graf, W. & Krueger, H. (1989). Ergonomic evaluation of user-interfaces by means of eye-movement data. In M. J. Smith, & G. Salvendy (Eds.), Work with Computers: Organizational, Management, Stress and Health Aspects (pp. 174-186). Amsterdam: Elsevier Science.

Graham, N. E. (1956). The speed and accuracy of reading horizontal, vertical, and circular scales. The Journal of Applied Psychology, 40(4), 282-232.

Harris, R. H. & Christhilf, D. M. (1980). What do pilots see in displays? Proceedings of the Human Factors Society 24th Annual Meeting, 22-26.

Harris, R. L. & Spady, A. A. (1985). Visual scanning behavior. Proceedings of the 1985 IEEE National Aerospace and Electronics Conference – NAECON 1985, 2, 85CH2189-9, 1032-1039.

Harris, R. L. Glover, B. J. & Spady, A.A. (1986). Analytical techniques of pilot scanning behavior and their application. (Technical Paper No. 2525). Hampton, VA: National Aeronautics and Space Administration, Langley Research Center.

Harris, R. L. Tole, J. R., Eprath, A. R. and Stephens, A. (1982). How a new instrument affects pilots' mental workload. Proceedings of the Human Factors Society 26th Annual Meeting, 1010-1013.

Hendrickson, J. L. (1988). Natural language human-computer dialogue: Menu-based natural language and visual performance. (Doctoral dissertation, Old Dominion University, Norfolk, VA).

Holahan, C. J., Culler, R. E., & Wilcox, B. L. (1978). Effects of visual distraction on reaction time in a simulated traffic environment. Human Factors, 20(4), 409-413.

Jones, D. H. (1985). An error-dependent model of instrument-scanning behavior in comercial airline pilots. (NASA Report No. CR-3908). Hampton, VA: National

Aeronautics and Space Administration, Langley Research Center.

Jonides, J. (1983). Further toward a model of the mind's eye's movement. Bulletin of the Psychonomic Society, 21(4), 247-250.

Just, M. A. & Carpenter, P. A. (1976). Eye Fixations and Cognitive Processes. Cognitive Psychology, 8, 441-480.

Kolers, P. A., Duchinsky, R. L. & Furguson, D. C. (1981). Eye movement measurement of readability of CRT displays. Human Factors, 23, 517-527.

Krappman, P. (1995). Distribution of visual attention revealed by eye movements and neural networks. In J.M. Findlay (Ed.), Eye Movement Research (pp. 514-525). Amsterdam: Elsevier Science.

Kundel, C. F. & Nodine, H. L. (1978). Studies of eye movements and visual search in radiology. In J. W. Senders , D. F. Fisher & R. A. Monty (Eds.), Eye Movements and the Higher Psychological Functions (pp. 317-327). Hillsdale, NJ: Lawrence Erlbaum.

Kundel, C. F., Nodine, C. F. & Toto, L. (1984). Eye movements and the detection of lung tumors in chest images. In A.G. Gale, & F. Johnson (Eds.), Theoretical and Applied Aspects of Eye Movements (pp. 415-427). North-Holland: Elsevier.

Landis, D., Slivka, R. M. & Jones, J. M. (1967). Evaluation of large-scale visual displays. (NTIS Report No. AD 651 372). Griffis Air Force Base, NY: Rome Air Development Center.

Lansdale, M. W., Jones, M. R., & Jones, M. A. (1989). Visual search in iconic and visual interfaces. Proceedings of the Ergonomic Society's Annual Conference. London: Taylor & Francis.

Lass, U., Klettke, W., Luer, G. & Ruhlender, P. (1991). Does thinking aloud influence the structure of cognitive processes? In R. Schmid & S. Zambabeiri (Eds.), Ouculometer Control and Cognitive Processes: Normal and Pathological Processes. Amsterdam: North-Holland.

Latimer, C. R. (1988). Eye-movement data: Cumulative fixation time and cluster analysis. Behavioral Research Methods, Instruments and Computations, 20(5), 437-470.

Lee, E. & MacGregor, J. (1985). Minimizing user search time in menu retrieval systems. Human Factors, 27(2), 157-162.

Levy-Schoen, A. (1981). Flexible and/or rigid control of oculometric scanning behavior. In D.F. Fisher, R.A. Monty, & J.W. Senders (Eds.), Eye Movements: Cognition and Visual Perception (pp. 299-314). Hillsdale, NJ: Erlbaum.

Luer, G., Hubner, R. & Lass, U. (1985). Sequences of eye movements in a problem solving situation. In R. Groner, G. W. McConkie & C. Menz (Eds.), Eye Movement and Human Information Processing. Amsterdam: North Holland.

Luer, G., Lass, U., Ulrich, M. & Schroiff, H. W. (1986). Changes in eye movement behavior in complex problem solving. In F. Klix & H. Hagendorf (Eds.), Human Memory and Cognitive Capabilities (pp. 197-209). Amsterdam: North Holland.

Mackworth, N. H. & Morandi, A. (1967). The gaze selects informative details within pictures. Perception and Psychophysics, 2, 547-552.

Moray, N. (1967). Where is attention limited? A survey and model. Acta Psychologica, 27, 84-92.

Moray, N. (Ed.), (1979). Mental Workload: Its theory and measurement. New York: Plenum Press.

Moray, N. & Rotenberg, I. (1989). Fault management in process control: Eye movements and action. Ergonomics, 32, 1319-1342.

Nelson, T. (1965). A file structure for the complex, the changing, and the indeterminate. Proceedings of the ACM National Conference, 84-100.

Nodine, C. F. & Kundel, H. L. (1982). Hidden figures, eye movements, and the perception of art. In R. Groner, & P. Fraisse (Eds.), Cognition and Eye Movements (pp. 345-357). North-Holland: Elsevier.

Norman, D. A. (1968). Towards a theory of memory and attention. Psychological Review, 75, 522-536.

Posner, M. I., Snyder, C. R. R. & Davidson, B. J. (1980) Attention and the detection of signals. Journal of Experimental Psychology, 109, 160-174.

Russo, J. & Rosen, L. (1975). An eye fixation analysis of multialternative choice. Memory and Cognition, 3, 267-276.

Russo, J. E. (1978). Adaptation of cognitive processes to the eye movement system. In J.W. Senders, D. F. Fisher & R. A. Monty (Eds.), Eye Movements and the Higher Psychological Functions (pp. 89-109). Hillsdale, NJ: Lawrence Erlbaum.

Sanders, M. S., & McCormick, E. J. (1993). Human factors in engineering and design. New York: McGraw Hill.

Scinto, L. F. & Barnette, B. D. (1986). An algorithm for determining clusters, pairs or singletons in eye-movement scan-path records. Behavioral Research Methods, Instruments and Computations, 18(1), 41-44.

Scott, D. (1992). Peripheral vision, conspicuity and VDU's. In E. J. Lovesey (Ed.), Contemporary Ergonomics. (pp. 41-46). London: Taylor and Francis.

Scott, D., & Findlay, J. M. (1991). Optimum display arrangements for presenting status information. International Journal of Man-Machine Studies, 35, 399-407.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom system. IEEE Transactions on Human Factors in Electronics, HFE-5, 1-6.

Senders, J. W. (1966). A re-analysis of the pilot eye movement data. IEEE Transactions of Human Factors in Electronics, HFE-7(2), 103-106.

Senders, J. W. (1983). Visual scanning processes. Tilburg, Netherlands: Tilburg University Press.

Senders, J. W., Elkind, J., Grignetti, M. C. & Smallwood, R. P. (1964). An investigation of the visual sampling behavior of human observers. (Technical Report No. NASA-CR-434).

Shannon, C. E. (1948). A mathematical theory of communication. Bell Systems Technical Journal, 27, 379-423.

Shapiro, K. L. & Raymond, J. E. (1989). Training of efficient oculomotor strategies enhances skill acquisition. Acta Psychologica, 71, 217-242.

Sheridan, T. B. (1970). On how often the supervisor should sample. IEEE Transactions on Systems, Science and Cybernetics, SSC-6, 140-145.

Shneiderman, B. (1987). Designing the user interface: Strategies for effective human-computer interaction. Reading, MA: Addison-Wesley.

Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. IEEE Computer, 16(8), 57-69.

Spady, A. A. (1978). Pilot scan patterns during simulated ILS approaches. (Technical Paper, No.1250), NASA.

Spady, A. A. Jones, D. H., Coates, G. D. & Kirby, R. H. (1982). The effectiveness of using real-time eye scanning information for pilot training. Proceeding of the Human Factors Society 26th Annual Meeting, 1014-1017.

Stark, L. & Ellis, N (1971). Eye movements and visual perception. Scientific American, 224, 34-43.

Stark, L. & Ellis, S. R. (1981). Scanpaths revisted: Cognitive models direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), Eye movements: Cognition and visual perception (pp. 524-537). Hillsdale, NJ: Erlbaum.

Stein, W. & Wewerinke, P. H. (1983). Human display monitoring and failure detection: Control theoretic models and experiments. Automatica, 19, 711-718.

Sutherland, I. E. (1963). SketchPad: A Man-Machine Graphical Communication System. AFIPS Spring Joint Computer Conference, 23, 329-346.

Tannas, L. E. (1985). Flat panel displays and CRTs. New York: Van Nostrand Reinhold Company.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. Journal of Experimental Psychology: Human Perception and Performance, 8, 194-214.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology - General, 114, 285-310.

Tullis, T. S. (1983). The formatting of alphanumeric displays: A review and analysis. Human Factors, 25(6), 657-682.

Unema, P. & Rotting (1985). Differences in eye movements and mental workload between experienced and inexperienced motor-vehicle drivers. In R. Groner, G. W. McConkie & C. Menz (Eds.), Eye Movement and Human Information Processing (pp. 194-202). Amsterdam: North Holland.

Van Delft, J. H. (1987). The development of a response sequence: A new description of human sampling behavior with multiple independent sources of information. Proceedings of the Human Factors Society 31st Annual Meeting, 151-155.

Wickens, C. D. (1987). Attention. In P. Hancock (Ed.), Human Factors in Psychology, (pp. 29-80). Amsterdam: North-Holland.

Wickens, C. D. (1992). Engineering psychology and human performance. Champaign-Urbana, IL: Harpers-Collins.

Wickens, C. D., & Andre, A. D. (1990). Proximity compatibility and information display: Effects of color, space, and object brightness of information integration. Human Factors, 32, 61-67.

Williams, A. J. & Harris, R. L. (1985). Factors affecting dwell times on digital displays. (Technical Memorandum No.86406). NASA.

Winer, B. J., Brown, D. R. & Michaels, K.M. (1991). Statistical principles in experimental design (3rd. Ed.) New York: McGraw-Hill.

Wright, R. B. & Converse, S. A. (1992). Method bias and concurrent verbal protocol in software usability testing. Proceedings of the Human Factors Society 36th Annual Meeting, 1220-1224.

Yarbus, A. L. (1967). Eye movements and vision. New York: Plenum Press.

Young L. R. & Sheena, D. (1975). Survey of eye movement recording methods. Behavior Research Methods and Instrumentation. 7, 397-429.

# APPENDIX A

## Source of Variation Table (Trial Time)

Dependent Variable: TIME

| Source | Sum of Squares | df | Mean Square | F | * = p<.05<br>** =p<.01 |
|---|---|---|---|---|---|
| C | 1112.110 | 2 | 556.055 | 10.906 | ** |
| D | 188.034 | 2 | 94.017 | 3.355 | |
| N | 91.597 | 1 | 91.597 | 3.493 | |
| X | 134.623 | 1 | 134.623 | 2.284 | |
| C * D | 382.616 | 4 | 95.654 | 5.081 | ** |
| D * N | 176.008 | 2 | 88.004 | 7.079 | ** |
| D * X | 23.028 | 2 | 11.514 | .411 | |
| C * N | 85.516 | 2 | 42.758 | 3.420 | |
| C * X | 8.016 | 2 | 4.008 | .079 | |
| N * X | .133 | 1 | .133 | .005 | |
| C * D * N | 285.583 | 4 | 71.396 | 4.129 | ** |
| C * D * X | 177.053 | 4 | 44.263 | 2.351 | |
| D * N * X | 15.394 | 2 | 7.697 | .619 | |
| C * N * X | 1.494 | 2 | .747 | .060 | |
| C * D * N * X | 50.434 | 4 | 12.609 | .729 | |
| S(X) | 589.507 | 10 | 58.951 | | |
| C * S(X) | 1019.737 | 20 | 50.987 | | |
| D * S(X) | 560.378 | 20 | 28.019 | | |
| N * S(X) | 262.224 | 10 | 26.222 | | |
| C * D * S(X) | 752.978 | 40 | 18.824 | | |
| D * N * S(X) | 248.625 | 20 | 12.431 | | |
| C * N * S(X) | 250.016 | 20 | 12.501 | | |
| C * D * N * S(X) | 691.703 | 40 | 17.293 | | |

C = Task Complexity

D = Information Density

N = Navigation Aid

X = Experience

S = Subject

## APPENDIX B

## Source of Variation Table (Error Rate)

Dependent Variable: ERROR

| Source | Sum of Squares | df | Mean Square | F | * = p<.05<br>** =p<.01 |
|---|---|---|---|---|---|
| C | 1.433 | 2 | .716 | 17.039 | ** |
| D | 4.083E-02 | 2 | 2.042E-02 | .242 | |
| N | 2.707E-03 | 1 | 2.707E-03 | .018 | |
| X | 2.007E-02 | 1 | 2.007E-02 | .837 | |
| C * D | .404 | 4 | .101 | 1.047 | |
| D * N | .141 | 2 | 7.039E-02 | .740 | |
| D * X | .504 | 2 | .252 | 2.981 | |
| C * N | .413 | 2 | .207 | 1.413 | |
| C * X | .511 | 2 | .256 | 6.079 | ** |
| N * X | .217 | 1 | .217 | 1.439 | |
| C * D * N | .230 | 4 | 5.741E-02 | .775 | |
| C * D * X | .162 | 4 | 4.057E-02 | .420 | |
| D * N * X | .230 | 2 | .115 | 1.210 | |
| C * N * X | .110 | 2 | 5.518E-02 | .377 | |
| C * D * N * X | .824 | 4 | .206 | 2.783 | * |
| S(X) | .240 | 10 | 2.398E-02 | | |
| C * S(X) | .841 | 20 | 4.205E-02 | | |
| D * S(X) | 1.691 | 20 | 8.453E-02 | | |
| N * S(X) | 1.506 | 10 | .151 | | |
| C * D * S(X) | 3.863 | 40 | 9.658E-02 | | |
| D * N * S(X) | 1.902 | 20 | 9.510E-02 | | |
| C * N * S(X) | 2.924 | 20 | .146 | | |
| C * D * N * S(X) | 2.961 | 40 | 7.404E-02 | | |

C = Task Complexity

D = Information Density

N = Navigation Aid

X = Experience

S = Subject

# APPENDIX C

## Source of Variation Table (Dwell Time)

Dependent Variable: DWELTIME

| Source | Sum of Squares | df | Mean Square | F | * = p<.05<br>** =p<.01 |
|---|---|---|---|---|---|
| C | 364.559 | 2 | 182.280 | 9.970 | ** |
| D | 56.823 | 2 | 28.411 | 2.456 | |
| N | 22.271 | 1 | 22.271 | 2.539 | |
| X | 124.807 | 1 | 124.807 | 1.621 | |
| C * D | 75.178 | 4 | 18.795 | 1.835 | |
| D * N | 48.951 | 2 | 24.476 | 4.253 | * |
| D * X | 11.525 | 2 | 5.763 | .498 | |
| C * N | 32.095 | 2 | 16.047 | 3.388 | |
| C * X | 1.370 | 2 | .685 | .037 | |
| N * X | 4.866 | 1 | 4.866 | .555 | |
| C * D * N | 73.222 | 4 | 18.305 | 2.189 | |
| C * D * X | 73.638 | 4 | 18.409 | 1.797 | |
| D * N * X | .934 | 2 | .467 | .081 | |
| C * N * X | 4.462 | 2 | 2.231 | .471 | |
| C * D * N * X | 18.607 | 4 | 4.652 | .556 | |
| S(X) | 770.155 | 10 | 77.016 | | |
| C * S(X) | 365.642 | 20 | 18.282 | | |
| D * S(X) | 231.386 | 20 | 11.569 | | |
| N * S(X) | 87.731 | 10 | 8.773 | | |
| C * D * S(X) | 409.783 | 40 | 10.245 | | |
| D * N * S(X) | 115.109 | 20 | 5.755 | | |
| C * N * S(X) | 94.742 | 20 | 4.737 | | |
| C * D * N * S(X) | 334.569 | 40 | 8.364 | | |

C = Task Complexity

D = Information Density

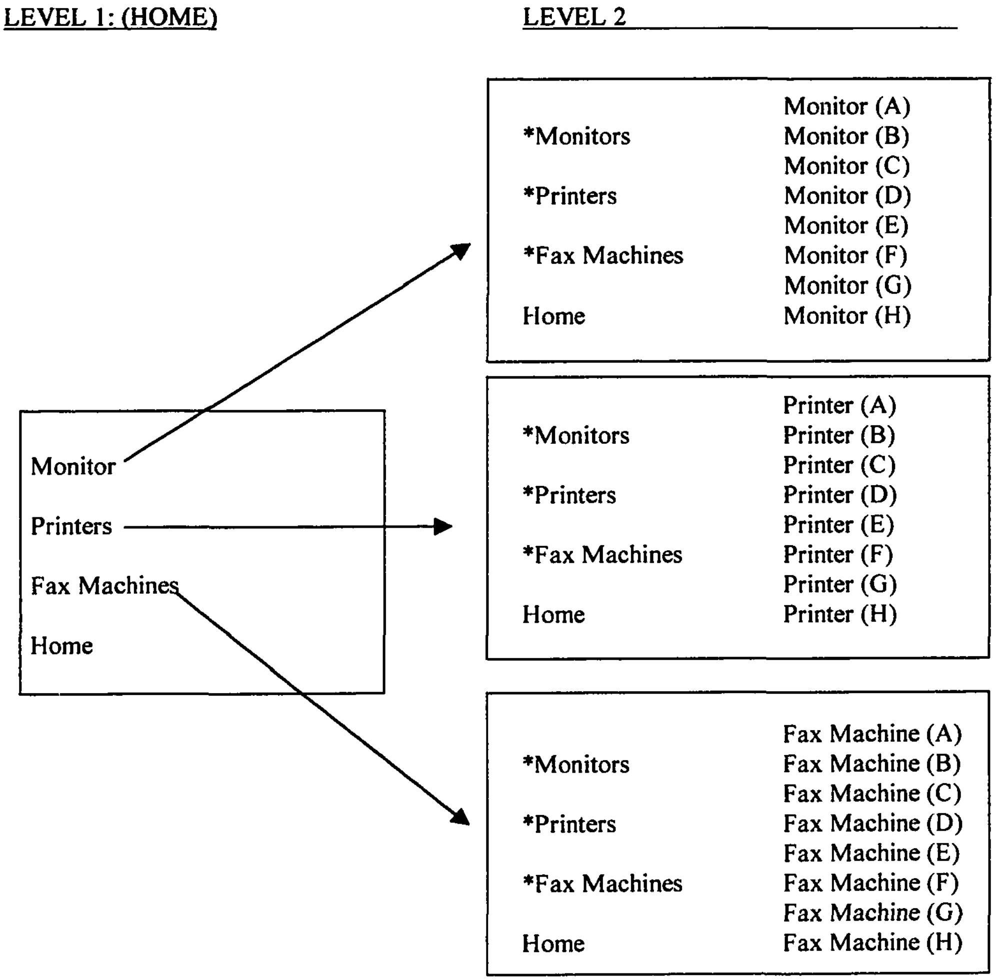
N = Navigation Aid

X = Experience

S = Subject

## APPENDIX D

## Overview of Screen Hierarchy



* Button not visible at level 2 in trials where the navigation aid is *absent.*

# APPENDIX E

## Task Descriptions

| Task Description | Task Complexity |
|---|---|
| Find the smallest monitor. | Low |
| Find the largest monitor. | Low |
| Find the least expensive monitor. | Low |
| Find the most expensive monitor. | Low |
| Find the smallest Samsung monitor. | Medium |
| Find the most expensive Samsung monitor. | Medium |
| Find the largest NEC monitor. | Medium |
| Find the least expensive NEC monitor. | Medium |
| Find the least expensive monitor between Sony Multiscan 15sx and ViewSonic 17EA. | High |
| Find the most expensive monitor between NEC MultiSync XV17+ and the Sony MultiScan 15sx. | High |
| Find the largest monitor between the Samsung SynchMaster 15GLe and MAG InnoVision DX 1595. | High |
| Find smallest monitor between the ViewSonic 17EA and the ViewSonic 15ES. | High |
| Find the printer with the smallest amount of memory. | Low |
| Find the printer with the largest amount of memory. | Low |
| Find the least expensive printer. | Low |
| Find the most expensive printer. | Low |
| Find the Lexmark printer with the smallest amount of memory. | Medium |
| Find the most expensive Lexmark printer. | Medium |
| Find the Epson printer with the largest amount of memory. | Medium |
| Find the least expensive Epson printer. | Medium |
| Find the least expensive printer between the Canon BJC 210 and the HP DeskJet 855Cse. | High |
| Find the most expensive printer between the Lexmark 1020 JetPrinter and Epson Stylus color IIs. | High |
| Find the printer with the largest amount of memory between the HP DeskJet 682C and the Canon BJC 610. | High |
| Find the printer with the smallest amount of memory between the Lexmark 2070 JetPrinter and the Canon BJC 210. | High |
| Find the fax machine with the smallest cost per page. | Low |
| Find the fax machine with the largest cost per page. | Low |
| Find the least expensive fax machine. | Low |
| Find the most expensive fax machine. | Low |

| | |
|---|---|
| Find the Panasonic fax machine with the largest cost per page. | Medium |
| Find the least expensive Panasonic fax machine. | Medium |
| Find the Brother fax machine with the smallest cost per page. | Medium |
| Find the most expensive Brother fax machine. | Medium |
| Find the least expensive fax machine between the Brother 625 and the Muratec M4500. | High |
| Find the most expensive fax machine between the HP OfficeJet 300 and the Sharp UX176. | High |
| Find the fax machine with the smallest cost per page between the Sharp UX176 and the Brother 825MC. | High |
| Find the fax machine with the largest cost per page between the Radio Shack TFX1032 and the Brother 625. | High |

# APPENDIX F

## Product Database

| Data Name | Data |
|---|---|
| MonitorName1 | NEC MultiSync XV15+ |
| MonitorName2 | Samsung SynchMaster 15GLe |
| MonitorName3 | Sony Multiscan 15sx |
| MonitorName4 | ViewSonic 15ES |
| MonitorName5 | MAG InnoVision DX1595 |
| MonitorName6 | Samsung SynchMaster 6Ne |
| MonitorName7 | NEC MultiSync XV17+ |
| MonitorName8 | ViewSonic 17EA |
| MonitorVariable1A | 13.7 inch |
| MonitorVariable1B | 13.8 inch |
| MonitorVariable1C | 13.9 inch |
| MonitorVariable1D | 14.0 inch |
| MonitorVariable1E | 14.3 inch |
| MonitorVariable1F | 15.9 inch |
| MonitorVariable1G | 15.3 inch |
| MonitorVariable1H | 15.8 inch |
| MonitorVariable2A | $480 |
| MonitorVariable2B | $430 |
| MonitorVariable2C | $450 |
| MonitorVariable2D | $380 |
| MonitorVariable2E | $390 |
| MonitorVariable2F | $700 |
| MonitorVariable2G | $850 |
| MonitorVariable2H | $660 |
| PrinterName1 | Epson Stylus Color IIs |
| PrinterName2 | Canon BJC 210 |
| PrinterName3 | Epson Stylus Color II |
| PrinterName4 | HP DeskJet 682C |
| PrinterName5 | Canon BJC 610 |
| PrinterName6 | Lexmark 2070 JetPrinter |
| PrinterName7 | HP DeskJet 855Cse |
| PrinterName8 | Lexmark 1020 JetPrinter |
| PrinterVariable1A | 15K memory |
| PrinterVariable1B | 62K memory |
| PrinterVariable1C | 56K memory |
| PrinterVariable1D | 512K memory |
| PrinterVariable1E | 96K memory |
| PrinterVariable1F | 5K memory |
| PrinterVariable1G | 812K memory |

| | |
|---|---|
| PrinterVariable1H | 10K memory |
| PrinterVariable2A | $190 |
| PrinterVariable2B | $150 |
| PrinterVariable2C | $230 |
| PrinterVariable2D | $300 |
| PrinterVariable2E | $500 |
| PrinterVariable2F | $160 |
| PrinterVariable2G | $850 |
| PrinterVariable2H | $660 |
| FaxName1 | Brother 825MC |
| FaxName2 | Panasonic KX F750 |
| FaxName3 | Brother 625 |
| FaxName4 | Radio Shack TFX 1032 |
| FaxName5 | HP OfficeJet 300 |
| FaxName6 | Panasonic KX F1000 |
| FaxName7 | Muratec M4500 |
| FaxName8 | Sharp UX 176 |
| FaxVariable1A | 6¢/page |
| FaxVariable1B | 2¢/page |
| FaxVariable1C | 5¢/page |
| FaxVariable1D | 4¢/page |
| FaxVariable1E | 8¢/page |
| FaxVariable1F | 9¢/page |
| FaxVariable1G | 3¢/page |
| FaxVariable1H | 7¢/page |
| FaxVariable2A | $300 |
| FaxVariable2B | $450 |
| FaxVariable2C | $240 |
| FaxVariable2D | $380 |
| FaxVariable2E | $470 |
| FaxVariable2F | $320 |
| FaxVariable2G | $430 |
| FaxVariable2H | $660 |

# APPENDIX G

## Computer Experience Questionnaire

I have been using personal computers for:
[ ] less than one month
[ ] 1-6 months
[ ] 7 months to a year
[ ] 1 to three years
[ ] more than three years (number of years: ______)

In general how do you feel about using computers?
[ ] like
[ ] dislike
[ ] indifferent

Check all the operating systems you use or have used and indicate length of time spend with each:
[ ] Windows 3.1 Length of time used:_____________
[ ] Windows 95 Length of time used:_____________
[ ] Windows NT Length of time used:_____________
[ ] Macintosh OS Length of time used:_____________
[ ] Unix Length of time used:_____________
[ ] Other _______ Length of time used:_____________

In terms of using a computer, I consider myself a:
[ ] Novice
[ ] Intermediate
[ ] Expert

I regularly use the following types of software program(s) (check all that apply)
[ ] word processor
[ ] spreadsheet
[ ] personal finance
[ ] games
[ ] electronic mail
[ ] CAD program
[ ] World Wide Web browser
[ ] other ________________________________________

## APPENDIX H

### Verbatim Instructions

In the upper portion of the computer screen is a white rectangle that will contain a description of the item you will need find in order to complete each task. This task description will remain visible in rectangle throughout the testing procedure.

The database you will use contains four areas: Home, Monitors, Printers, and Fax Machines. Each trial will start at the Home screen. In order to navigate to another area of the database, click on the button for that area. The other three buttons are always visible from the Home screen but they are not always visible from other screens. You can always get back to the Home screen by clicking the Home button.

The task description will provide you with a description of an item in the database that you will need to find. All items described in the task descriptions exist in the database. When you find the target item, click on it with the cursor using the left mouse button. Following each correct response you will be shown a confirmation screen indicating that your response was correct and another task description will be displayed.

At this time I would like to guide you through 4 practice trials to make sure that you understand the procedure.

You will be presented with a total of 36 task descriptions. When asked to, please click the "Start" button and the first task description will appear. Again, you will be looking for an item in the database and clicking it with your mouse when you find it. Please respond to each of the 36 display sets as quickly as possible with no more than 5% errors.

I will be in the room during the procedure monitoring the eye-tracking equipment. The computer will tell you when the procedure is finished. Are there any questions regarding the procedure?

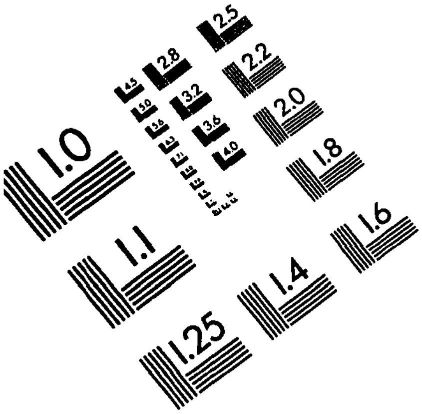
Please click the “Start” button to begin.

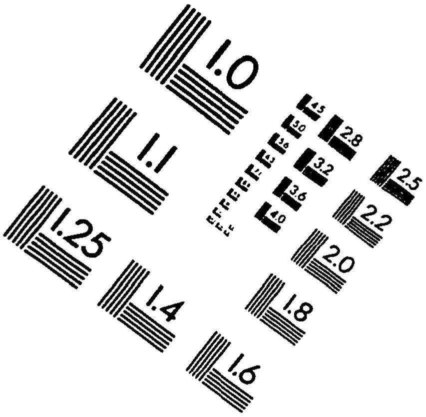
# VITA

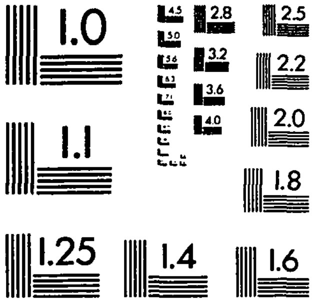
Orhan Beckman was born in Oconomowoc, Wisconsin on July 16, 1967. In May of 1990 he received his Bachelor of Arts in Psychology from Earlham College. In the fall of 1991 he entered the Graduate School of Old Dominion University. In May of 1994 he received the Master of Science degree in Experimental Psychology from Old Dominion University. Orhan performed his doctoral internship at the Vancouver Printer Division of Hewlett-Packard in Vancouver, Washington. In 1995 he joined the research and development staff at Hewlett-Packard as a Human Factors Engineer. Hewlett-Packard funded Orhan's studies at Old Dominion University for a year through the Resident Fellowship program. He subsequently received his Doctor of Philosophy degree in Industrial/Organizational Psychology in the fall of 1998.

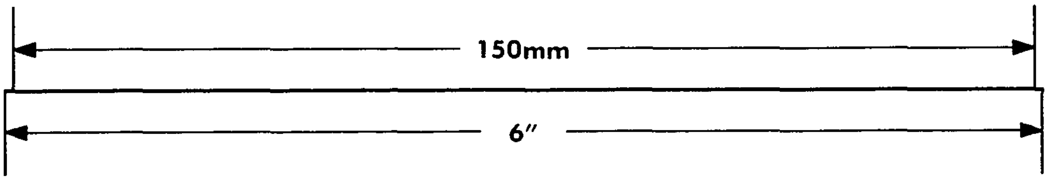
Department of Psychology
MGB Room 250
43rd and Hampton Boulevard
Norfolk, Virginia
23529

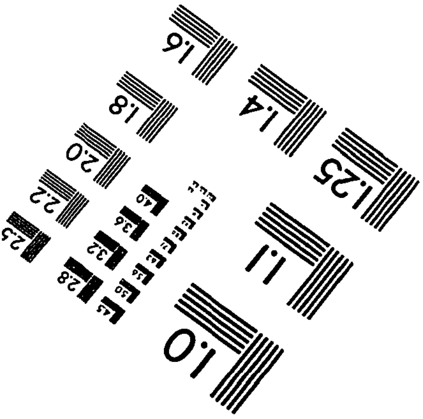
# IMAGE EVALUATION TEST TARGET (QA-3)











APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989